Structure Determination of Neurotensin Receptor 1 Variants Functionally Expressed in *E. coli*

DISSERTATION

zur

Erlangung der naturwissenschaftlichen Doktorwürde
(Dr. sc. nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der

Universität Zürich

von

Pascal Egloff

von

Niederohrdorf AG

Promotionskomitee

Prof. Dr. Andreas Plückthun (Vorsitz)

Prof. Dr. Stephan Grzesiek

Prof. Dr. Raimund Dutzler

Prof. Dr. Nenad Ban

Zürich 2014

Department of Biochemistry
University of Zurich
Winterthurerstrasse 190
8057 Zurich
Switzerland

To my friends and family

Abstract

In multicellular eukaryotes, G protein-coupled receptors (GPCRs) play important roles in intercellular communication and in the perception of environmental conditions. They bind to various classes of ligands, including hormones, a light-sensitive compound, pheromones and neurotransmitters. Upon binding, they activate signaling pathways, thus ultimately causing a cellular response. A better understanding of GPCR signal-transduction mechanisms would improve our knowledge about fundamental processes in eukaryotic cells and it would facilitate the development of new medicine.

Low expression levels and instability in solution have limited structural insights to very few selected members of this large protein family. It is particularly challenging to express GPCRs in prokaryotes, such as *E. coli* – a problem that renders many biophysical experiments impossible. Previously, researchers in the Plückthun laboratory developed directed evolution methods in *E. coli*, which generated neurotensin receptor 1 (NTR1) variants with significantly increased expression levels. In order to learn more about NTR1 and to prove the applicability of the evolutionary approach to structural biology, it was the goal of this doctoral work to determine the crystal structure of an evolved GPCR variant.

A novel and unique purification procedure in the GPCR field was developed, which is based on a disposable, high-affinity ligand column and allowed the isolation of more than 10 mg of functional receptors directly from solubilized *E. coli* cells (no membrane preparation). The economical and time-efficient strategy was the foundation for the vapor diffusion crystallization of four constructs of three different evolved NTR1 variants bound to the neurotensin agonist. Initial purification problems using standard IMAC procedures also inspired the development of further evolution methods, selecting directly for long-term stability in detergent micelles. The variant TM86V that yielded the highest resolution structure at 2.75 Å was functionally characterized. The crystallized constructs exhibited residual ligand-dependent signaling, internalization and wild-type-like agonist and antagonist affinities. The structures of the three different NTR1 variants TM86V, OGG7 and HTGH4 are fully consistent with all biochemically defined ligand-contacting residues, thus underlining the applicability of directed evolution technologies to GPCR structural biology.

All structures are highly similar, but they exhibit major differences to another NTR1 structure of a signaling-deficient variant that was fused to T4 lysozyme for structure determination (PDB-ID: 4GRV). The differences are observed in the ligand-binding pocket and at the cytosolic region, where the amphipathic helix 8 in the signaling-deficient NTR1 variant is surprisingly absent. An analysis of helix 8 stability determinants between NTR1 and other crystallized GPCRs suggests that the occupancy of the canonical position of the amphipathic

helix is reduced to various extents in many receptors, and we have elucidated the sequence determinants for a stable helix 8. Our analysis also provides a structural rationale for the long-known effects of C-terminal palmitoylation reactions on GPCR signaling, receptor maturation and desensitization.

In the last part of this dissertation, the purification methods were applied to obtain the first NMR spectra of isotope-labeled, fully deuterated and functionally expressed GPCRs. Most backbone amide resonances could be observed in a ^{1}H-^{15}N TROSY spectrum after detergent optimization and the spectral quality implied that assignments are possible, suggesting that unique data on the dynamics of a model GPCR might be obtained in the near future.

Zusammenfassung

In mehrzelligen Eukaryoten übernehmen G-Protein-gekoppelte Rezeptoren (GPCRs) Kernaufgaben der interzellulären Kommunikation und bei der Wahrnehmung von Signalen der Umwelt. Sie binden sehr unterschiedlichen Typen von Liganden, zum Beispiel Hormone, Chromophore, Pheromone und Neurotransmitter. Die extrazelluläre Liganden-Bindung führt zur Aktivierung von intrazellulären Signalkaskaden und schlussendlich zu einer zellulären Reaktion. Mehr Erkenntnisse über den Signal-Übertragungs-Mechanismus der GPCRs würden unser Verständnis fundamentaler Prozesse in eukaryotischen Zellen verbessern und die Entwicklung neuer Medikamente vereinfachen.

Tiefe Expressionsniveaus und Instabilität in Lösung haben dazu geführt, dass nur wenige GPCRs strukturell untersucht werden konnten. Im Besonderen ist es schwierig diese Rezeptoren von Prokaryoten, wie zum Beispiel von *E. coli*, herstellen zu lassen, was einige biophysikalische Experimente nahezu verunmöglicht. Vor Beginn dieser Doktorarbeit haben Wissenschaftler des Plückthun-Labors sogenannte „zielgerichtete Evolutionsmethoden“ (directed evolution) in *E. coli* entwickelt, die es erlaubten, besser herstellbare Neurotensinrezeptor 1 (NTR1) Varianten zu generieren. Um neue strukturelle Erkenntnisse über NTR1 zu generieren und um die Anwendbarkeit der gerichteten Evolutionsprinzipien zu beweisen, war es das Ziel dieser Dissertation, eine hochaufgelöste Kristallstruktur einer evolvierten Rezeptorvariante zu ermitteln.

Basierend auf einer Einweg-Ligandensäule, wurde eine bisher einzigartige Reinigungsmethode im GPCR-Forschungsfeld entwickelt, die die Anreicherung von mehr als 10 mg funktionalem Rezeptor direkt aus aufgelösten *E. coli* Zellen ohne Membranpreparation erlaubt. Diese ökonomische und Zeit-effiziente Strategie war die Grundlage für die Kristallisation von vier Agonist-gebundenen Rezeptorkonstrukten dreier unterschiedlicher NTR1 Varianten. Anfängliche Reinigungsprobleme mit Standard-Methoden haben zudem die Entwicklung alternativer Evolutionsmethoden inspiriert, die direkt auf dauerhafte Rezeptorstabilität in gelöster Form selektieren. Die NTR1 Variante TM86V ergab die höchstaufgelöste Struktur (2.75 Å) und wurde deshalb funktionell charakterisiert. Die kristallisierten Konstrukte zeigten trotz Mutationen eine verbleibende Liganden-abhängige Signalübertragung, Internalisation und Wildtyp-ähnliche Agonist- und Antagonistaffinität. Die Strukturen der drei unterschiedlichen NTR1 Varianten TM86V, OGG7 und HTGH4 entsprechen allen biochemisch definierten Ligandkontakten und belegen deshalb die Anwendbarkeit der gerichteten Evolution für die Strukturbestimmung von GPCRs.

Alle unsere Strukturen sind sich sehr ähnlich, aber sie unterscheiden sich stark von einer anderen NTR1 Struktur, die mithilfe einer Lysozymefusion bestimmt wurde, welche keine

Signalübertragung zulässt (PDB-ID: 4GRV). Die Unterschiede sind in der Ligandenbindungstasche und beim intrazellulären Bereich, wo zum Beispiel die ganze amphipathische Helix 8 entweder existiert oder wie beim nicht-funktionalem Konstrukt überaschenderweise abwesend ist. Eine Analyse der strukturellen Einflussgrössen bezüglich Helix 8-Stabilität ergab, dass die Besetzung („occupancy“) der typischen Helix 8 Position in verschiedenen GPCRs unterschiedlich stark ist. Des Weiteren erklärt sie, wie die bekannten Effekte der C-terminalen GPCR-Palmitoylierungsreaktionen auf die Signalübertragung, die Rezeptormaturierung und Desensitierung entstehen könnten.

Als letzter Teil dieser Arbeit wurde die entwickelte Reinigungsmethode eingesetzt, um die ersten NMR Spektren voll deuterierter, isotopenmarkierter und funktional hergestellter GPCRs zu messen. Die meisten Amid-Resonanzen des Proteinrückgrats sind in einem ^{1}H-^{15}N TROSY sichtbar nach Detergenzoptimierungen. Dies deutet darauf hin, dass eine Zuordung der meisten Resonanzen zu spezifischen Aminosäuren möglich ist und neue Einsichten bezüglich GPCR-Dynamik nun zugänglich werden könnten.

Content in brief

Chapter 1

Introduction: Evolution, GPCRs and Membrane Protein Technologies **1**

Chapter 2

A Cleavable Ligand Column for the Rapid Isolation of Large Quantities of Homogeneous and Functional Neurotensin Receptor 1 Variants from *E. coli*. **29**

Chapter 3

Structure of Signaling-Competent Neurotensin Receptor 1 Obtained by Directed Evolution in *Escherichia coli*. **53**

Chapter 4

Discussion and Outlook – First Experiments Towards GPCR Dynamics **75**

Chapter 5

Appendices: manuscript, abbreviations, expression plasmid, generalizability of purification principle, curriculum vitae, acknowledgments. **93**

Contents

Chapter 1 – Introduction 1
1.1. Historical aspects 2
1.2. G protein-coupled receptors 2
1.2.1. Structure 2
1.2.1. Function 3
1.2.1. Activation mechanism 5
1.2.1. Neurotensin receptor 1 7
1.3. Membrane protein structure determination 8
1.3.1. Alternative membrane mimics and the special case of GPCRs 9
1.3.2. Lipidic cubic phase crystallization 10
1.3.3. Screening the diversity of life and alanine-scanning 12
1.3.4. Directed evolution for structural studies of GPCRs 14
1.4. Goal of this work 17
1.5. Initial progress 18
1.5.1. Previous work 18
1.5.2. Implementation of the “Grisshammer protocol” 18
1.5.3. An alternative purification procedure 20
1.5.1. Conclusions and further strategy development 21
1.6. Thesis outline 22
1.7. References 23
Chapter 2 29
2.1. Manuscript 31
2.2. Abstract 32
2.3. Introduction 33
2.4. Materials and Methods 34
2.4.1. Materials 34
2.4.2. Construct design 34
2.4.3. Expression and purification of pD-NT contructs 35
2.4.4. Coupling of pD-NT constructs to NHS-activated Sepharose 36
2.4.5. Large-scale expression and purification of agonist-bound NTR1 variants 36
2.4.6. Identification of mutant pD-NT columns for antagonist-bound receptor purifications 38
2.4.7. Large-Scale antagonist-bound TM86V purification 38
2.5. Results and Discussion 39

2.5.1. Purification strategy and ligand column design 39
2.5.2. Ligand production and coupling efficiency 40
2.5.3. Large-Scale preparation of functional NTR1 variants 41
2.5.4. Purification of antagonist-bound TM86V 44
2.5.5. Generalizability of purification principle 46
2.6. Conclusions 47
2.7. References 48
2.8. Supplementary information 50
2.9. Supplementary reference 52
Chapter 3 53
3.1. Published article 54
3.2. Supporting information 62
Chapter 4 75
4.1. The potential of crystallized NTR1 variants for GPCR research 76
4.2. NMR 78
4.2.1. Expression in minimal medium and purification considerations 79
4.2.2. Detergent and temperature optimization using ^{1}H- ^{15}N TROSY experiments 80
4.2.1. Evidence of changing receptor dynamics upon binding of different ligands 83
4.3. Outlook and research suggestions 84
4.3.1. Nanodisc reconstitution 84
4.3.1. NMR assignment 86
4.3.2. Generating additional insights into NTR1 function by crystal structures 86
4.3.3. Suggestion: Engineering of mutants with specific signaling characteristics 86
4.4. Materials and Methods 88
4.4.1. Expression in minimal medium (adapted Wagner-lab protocol) 88
4.4.2. NMR experiments on TM86V-ΔIC3 and HTGH4-ΔIC3 89
4.4.3. Circular dichroism (CD) spectroscopy 89
4.4.4. Nanodisc reconstitution 89
4.5. References 91
Chapter 5 93
5.1. Manuscript 95
5.1.1. Abstract 96
5.1.2. Introduction 98
5.1.3. Materials and Methods 99
5.1.3.1 Stabilisation of NTS1 using CHESS 99
5.1.3.2 Stability analysis of 96 individual NTS1 variants from the selected population 100

5.1.3.3 Construct optimization ... 101
5.1.3.4 Stability comparison of engineered NTS1 variants ... 101
5.1.3.5 KingFisher saturation binding assays ... 102
5.1.3.6 KingFisher competition binding assays ... 103
5.1.4. Results ... 103
5.1.4.1 CHESS based evolution of apo-state stable NTS1 variants ... 103
5.1.4.2 Isolating NTS1-H4 from the CHESS selected population ... 103
5.1.4.3 Comparison of the stability of NTS1-H4 to other NTS1 variants ... 104
5.1.4.4 Saturation binding of NT to solubilised NTS1-H4 ... 104
5.1.4.5 Competition binding assays using solubilised NTS1-H4 ... 105
5.1.4.6 Sequence of NTS1-H4 ... 105
5.1.5. Discussion ... 105
5.1.6. References ... 115
5.2. Abbreviations ... 118
5.3. Plasmid for large-scale expression of NTR1 variants ... 120
5.4. Potential applicability of purification principles for other GPCRs ... 121
5.5. References ... 122
5.6. Curriculum vitae ... 123
5.7. Acknowledgments ... 125

Contribution statements

Chapter 1

I performed the following experiments: NTR1-D03 expression screens using various expression conditions and several *E. coli* strains. Solubilization tests with NTR1-D03. Purification trials according to the protocol of the Grisshammer group. Purification trials by standard "IMAC approaches" (according to the group of Raimund Dutzler). Radioligand-binding assays in solution to monitor receptor activity in the presence and absence of ligand. I wrote the chapter and created all illustrations (figures 1.3C, 1.6 and 1.7 were adapted from the indicated sources).

Chapter 2

I carried out the following theoretical work: Literature research to identify the most suitable ligand production and ligand coupling procedure. Design of the pD-NT construct and of the receptor expression constructs. Development of the new NTR1 purification strategy, which yielded NT-bound receptor. Design of the new strategy for the purification of antagonist-bound receptor.

I performed the following experiments: Molecular cloning of pD-NT constructs and optimization of their expression conditions in a 50 L fermenter. Purification of pD-NT constructs. Establishment and optimization of small- and large-scale coupling procedures. Molecular cloning of NTR1 expression constructs. Expression of receptor variants in *E. coli*. Purification of NTR1 variants. Stability analyses and absorbance measurements to characterize antagonist-binding. I wrote the entire manuscript.

Chapter 3

I carried out the following experiments: Radioligand-binding assays for the determination of binding affinities. Molecular cloning of receptor constructs with loop deletions and shortened termini. Purification of NTR1 variants according to the developed protocol. Crystallization of NTR1 variants. Crystal harvesting, cryoprotection and freezing in liquid propane or liquid nitrogen. Crystal mounting and data collection at the Swiss Light Source (Villigen).

I performed the following theoretical work: Diffraction data analysis and processing (indexing, integration, scaling). Molecular replacement and refinement. Structural analyses and generation of illustrative figures of the structures. I wrote the entire manuscript.

Chapter 4

The following experiments were carried by Dr. Franz Hagn and myself as a team: Purification of NTR1 variants in various detergents for 2D-TROSY tests. TROSY measurements of NTR1 variants in n-dodecyl- and n-decyl-β-D-maltopyranoside. Nanodisc reconstitution optimization for NTR1 variants.

I performed the following theoretical work: Structural analyses and subsequent generation of figures 4.1A and 4.1B. Analysis of 454-sequencing data and generation of figure 4.1C. Adaption of figures 4.2 – 4.5 from the indicated sources. I wrote the entire chapter.

Chapter 5

My receptor purification experiences contributed in many discussions with Dr. Daniel J. Scott to the design of the novel selection strategy described in section 5.1. I generated the structure figures of the manuscript in chapter 5.1 and helped writing it. I constructed the plasmid described in section 5.3 and carried out the literature research relevant for section 5.4. I wrote sections 5.2 – 5.7.

Chapter 1 – Introduction

Evolution, GPCRs and Membrane Protein Technologies

Content

Chapter 1 – Introduction 1
1.1. Historical aspects 2
1.2. G protein-coupled receptors 2
1.2.1. Structure 2
1.2.1. Function 3
1.2.1. Activation mechanism 5
1.2.1. Neurotensin receptor 1 7
1.3. Membrane protein structure determination 8
1.3.1. Alternative membrane mimics and the special case of GPCRs 9
1.3.2. Lipidic cubic phase crystallization 10
1.3.3. Screening the diversity of life and alanine-scanning 12
1.3.4. Directed evolution for structural studies of GPCRs 14
1.4. Goal of this work 17
1.5. Initial progress 18
1.5.1. Previous work 18
1.5.2. Implementation of the "Grisshammer protocol" 18
1.5.3. An alternative purification procedure 20
1.5.1. Conclusions and further strategy development 21
1.6. Thesis outline 22
1.7. References 23

1.1. Historical aspects

From a human perspective, the diversity of life on earth can be fascinating. Approximately 8.7 million distinct eukaryotic species exist[1] and our perceptional limits are challenged by their variability in morphology, lifestyle and organization at the tissue and cellular levels. Charles Darwin offered in 1859 an explanation for the existence of diverse life forms and suggested that they arose by common descent from evolving populations due to natural selection[2]. After the modern evolutionary synthesis, which combined his concepts, genetics and a body of evidence from other fields (1936-1947), we have arrived in the age of molecular life sciences to date. Like Darwin, who observed his first signs of evolution during the long and hazardous voyage on HMS Beagle, many modern researchers are setting out to another difficult journey these days – the exploration of life at the subcellular scale. Surprisingly, Darwinian principles – once generating today's diversity – are applied now to explore the living at molecular levels.

1.2. G protein-coupled receptors

One of the best described aspects among the tree of life is the distinction between prokaryotic and eukaryotic cells. In contrast to bacteria and archaea, eukaryotic cells contain a nucleus and other organelles like the endoplasmatic reticulum or mitochondria. Apparently, they are also different at the level of the genome, since many "younger" protein families are only present in one or the other domain. G protein-coupled receptors (GPCRs), for example, which exist at the interface of a cell's external and internal environment, appear to be restricted to eukarya[3]. In this domain and in particular in metazoans, however, they have evolved to play major physiological roles: They mediate most cellular responses to neurotransmitters and hormones and at the same time, they are crucial for vision, olfaction and taste[4]. In humans, more than 3 % of the open reading frames encode GPCRs, which makes them to the largest membrane protein family in this species[5].

1.2.1. Structure

GPCRs consist of seven transmembrane-spanning α-helices that are separated by extra- and intracellular loop regions (Figure 1.1). In vertebrates, they are commonly divided into five classes based on their structural and sequence similarity: class A (rhodopsin), class B (secretin), class C (glutamate), adhesion and Frizzled/Taste2. Here, the focus will be on class A GPCRs, which is by far the largest and most diverse class in terms of ligands. All members of this class exhibit small intra- and extracellular regions, which can be highly variable on the sequence level. Some motifs in the transmembrane region are, however, conserved to some extent, implying shared structural features and activation mechanisms (Figure 1.1).

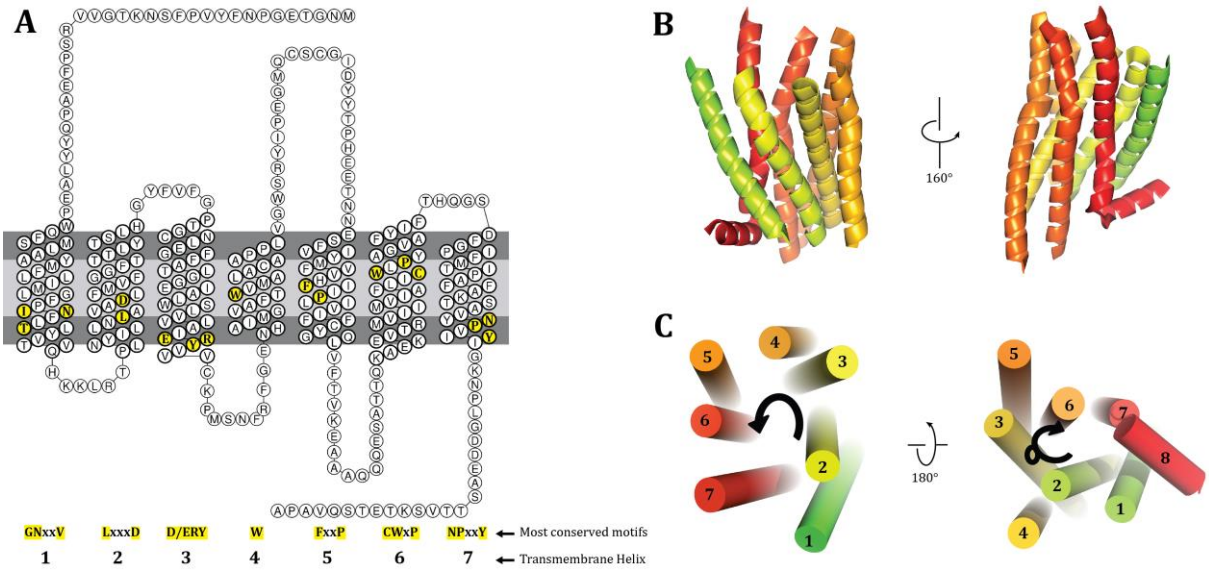


Figure 1.1 Overall architecture of class A GPCRs. A, Snake-plot of bovine rhodopsin, which was the first GPCR yielding high resolution crystal structures[6]. The most conserved motifs are depicted in yellow. Note that the termini, the loops and also parts of the transmembrane helices (TMs) are highly variable among class A GPCRs. As evident from the illustration, rhodopsin exhibits not all of the most conserved motifs (see TM1), which exemplifies the low sequence conservation of GPCRs. Extracellular parts are frequently glycosylated (not shown) and many GPCRs encode one or more palmitoylation sites (cysteines) in close proximity to the amphipathic H8 near the C-terminus (not the case for rhodopsin). B, view parallel to the membrane plane on the GPCR helical bundle in a prototypical inactive conformation (loops and termini are omitted). TMs are colored from N-terminus to C-terminus (green to red). C, Extracellular (left) and intracellular (right) view onto the transmembrane bundle. Note that class A GPCRs exhibit a circular and almost sequential arrangement of TMs. The main irregularity regarding sequential TM stacking are the exchanged intracellular endings of TM3 and 4, which is due to a pronounced tilt of TM3 relative to the other helices.

1.2.1. Function

The classical function of GPCRs is the coupling of an extracellular signal to the activation of heterotrimeric G proteins, which leads to the modulation of downstream effector proteins. Taking neurotensin receptor 1 (NTR1) as an example, the binding of neurotensin (or the related neuromedin N peptide) in the extracellular half of the TM bundle, leads to guanine-nucleotide exchange factor (GEF) activity at the GDP-bound heterotrimeric G protein[7]. After nucleotide exchange to GTP, the heterotrimeric G protein dissociates into α- and βγ-subunits, which in turn stimulate or inactivate the production of second messengers at downstream signaling partners[8].

At least 16 different G protein α subunits exist[9 10], which can be categorized into either $G\alpha_s$ $G\alpha_q$, $G\alpha_i$ or $G\alpha_{12}$. In addition five β and eleven γ subunits[11; 12] are encoded in mammals and

together, the three subunits form trimeric complexes of various stabilities. GPCRs exhibit preferential coupling to one or several heterotrimeric G proteins and vice versa. NTR1, for example, was described to couple mainly via G_q, which triggers production of inositol 1,4,5-trisphosphate (IP3) by the activation of phospholipase C (Figure 1.2A)[13], but also via G_s and G_i that modulate the activity of the adenylate cyclase and hence the level of cAMP in the cytosol (Figure 1.2A)[14-16]. The precise signaling behavior of each GPCR is likely dependent on the physiological context – e.g. on the expression levels of alternative signaling partners.

In humans, more than 600 different class A GPCRs are encoded[17] and expressed depending on the cell type, developmental stage and additional factors. The various G protein subunits can potentially combine into several hundred heterotrimers with various coupling properties in terms of GPCR preference and downstream effectors. Many heterotrimers have already been detected and were proven to couple to GPCRs (personal communication, M. Hillenbrand). This diverse set of signaling modules may represent a key adaptation of multicellular eukaryotes, as it allowed fine-tuned intercellular communication and thus a complex, well-balanced interplay of highly specialized tissues and organs[18].

An important feature of GPCRs is their waning responsiveness to repeated stimulation with agonists, a process termed desensitization. One part of this process is the internalization of activated receptors by endocytosis, which was also shown to occur for NTR1 upon agonist binding (Figure 1.2B)[19; 20]. Desensitization in general is usually initiated by GPCR kinases that specifically phosphorylate residues at activated GPCRs (mostly serines and threonines in proximity to the C-terminus after helix 8). Phosphorylated, agonist-bound receptors are then recognized by β-arrestins, which prevent G protein binding and thus block this type of signaling[8]. In recent years, it became apparent that β-arrestin-binding is the key trigger of internalization, as it initiates clathrin-mediated endocytosis, followed by either trafficking the receptor-β-arrestin complex to the lysosome or by receptor recycling to the cell membrane. Interestingly, β-arrestin can also signal via alternative "non-classical" pathways, for example via tyrosine or mitogen-activated protein (MAP) kinases[8; 21; 22]. Many GPCRs can exhibit preferential signaling, either towards the classical, G protein-mediated route or via β-arrestin. On one hand, the "choice" of signaling pathways depends on the particular kind of ligand bound to the receptor – a broadly accepted concept termed "biased signaling"[23] – and on the other hand, it depends on the expression levels of the signaling partners, making the response cell-type specific.

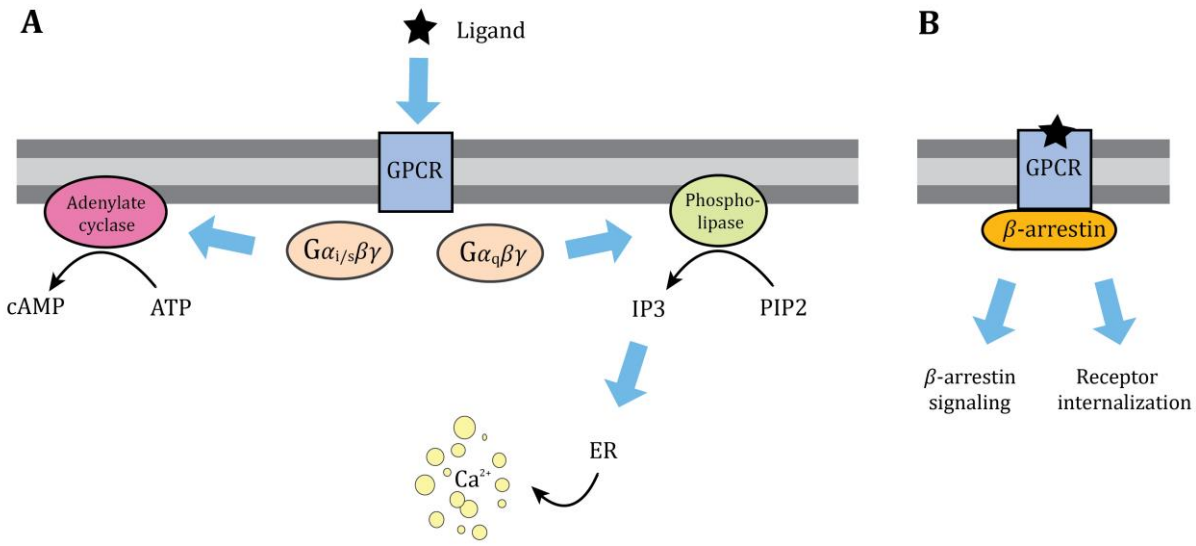


Figure 1.2 Overview on the most investigated signaling pathways of class A GPCRs. A, $G\alpha_s$ and $G\alpha_i$ regulate the cyclic AMP production, whereas $G\alpha_q$ activates phospholipase C to produce inositol 1,4,5-trisphosphate (IP3) upon hydrolysis of phosphatidylinositol 4,5-bisphosphate (PIP2). B, Receptor desensitization/internalization and alternative signaling behaviors are mediated by β-arrestins. Four types of β-arrestins are encoded in the human genome: Arrestin1 and arrestin4 are expressed in retinal rods and cones, respectively, whereas arrestin2 (also called β-arrestin1) and arrestin3 (also called β-arrestin2) are ubiquitously distributed and interact with most GPCRs[8].

1.2.1. Activation mechanism

As described above, at least 3 different types of proteins interact with the intracellular regions of activated GPCRs: Heterotrimeric G proteins, GRKs and β-arrestins. To date, mainly antagonist- or inverse agonist-bound GPCR structures are available, which represent inactive states that cannot bind to any signaling partner at the cytosolic side. Two types of high resolution data sets on GPCRs in complex with intracellular signaling partners were however obtained recently: First, rhodopsin bound to the C-terminal helix of the α-subunit of transducing (ocular heterotrimeric G protein)[24-28], and second, the β_2-adrenergic receptor (β_2AR) in complex with the heterotrimeric G_s protein[20]. Comparing the inactive-state structures of both receptors with their respective structures in the G protein complex, it can be observed that in both cases TM6 – and to a smaller extent TM5 – shift outward at the intracellular side upon G protein binding. The center of the "open" cytosolic receptor region is occupied by the C-terminal helix of the α-subunit of the G protein, which represents the main interaction site of the two signaling partners (Figure 1.3A and B). As rhodopsin is not a typical GPCR (highly expressed in certain tissues, highly thermostable, no diffusible ligand, activated by light), the following discussion about GPCR activation will focus on β_2AR.

Numerous structures of β_2AR were determined over the last 6 years and two interesting structural themes were observed: First, in the absence of G protein or G protein mimics, all

crystal structures of β_2AR showed "closed" conformations at the intracellular side, which appeared to be incompatible with G protein binding. This came as a surprise, since the receptors were crystallized in different space groups, by different crystallographic techniques and in complex with different ligands of alternative pharmacologies (inverse agonists, antagonists and agonists). Second, the differences at the extracellular side between inverse agonist-bound structures and ternary complexes (agonist/β_2AR/G protein) are extremely small (Figure 1.3A). How does signaling work, if only minor structural changes occur in the ligand-binding pocket and no active states can be observed at the G protein interacting surface?

The Kobilka group, which determined most of the β_2AR structures, proposed that the receptor exhibits a conformational ensemble of several distinct inactive and active states[29]. Using NMR[30; 31] and double electron-electron transfer (DEER, personal communication, Aashish Manglik), they provided evidence for a destabilization of inactive states upon agonist binding. They further suggest that active states are not stabilized by agonist-binding, but only upon interaction with G protein. Unlike the highly specialized rhodopsin, which appears to be an efficient two-state, light-activated switch between relatively well defined inactive and active states[32], β_2AR may be a more versatile, albeit less efficient molecular machine. It appears to couple ligand-binding to several different signaling processes inside the cell by sampling an ensemble of conformations with alternative binding-specificities – a behavior that might be prototypical for GPCRs.

An improved understanding of GPCR function will require a more detailed assessment of conformational dynamics, measurements of time-scales from milliseconds to seconds, computational efforts to rationalize these observations and the determination of further crystal structures of alternative receptors in complex with many cytosolic interaction partners, such as alternative G proteins, GRKs and β-arrestins.

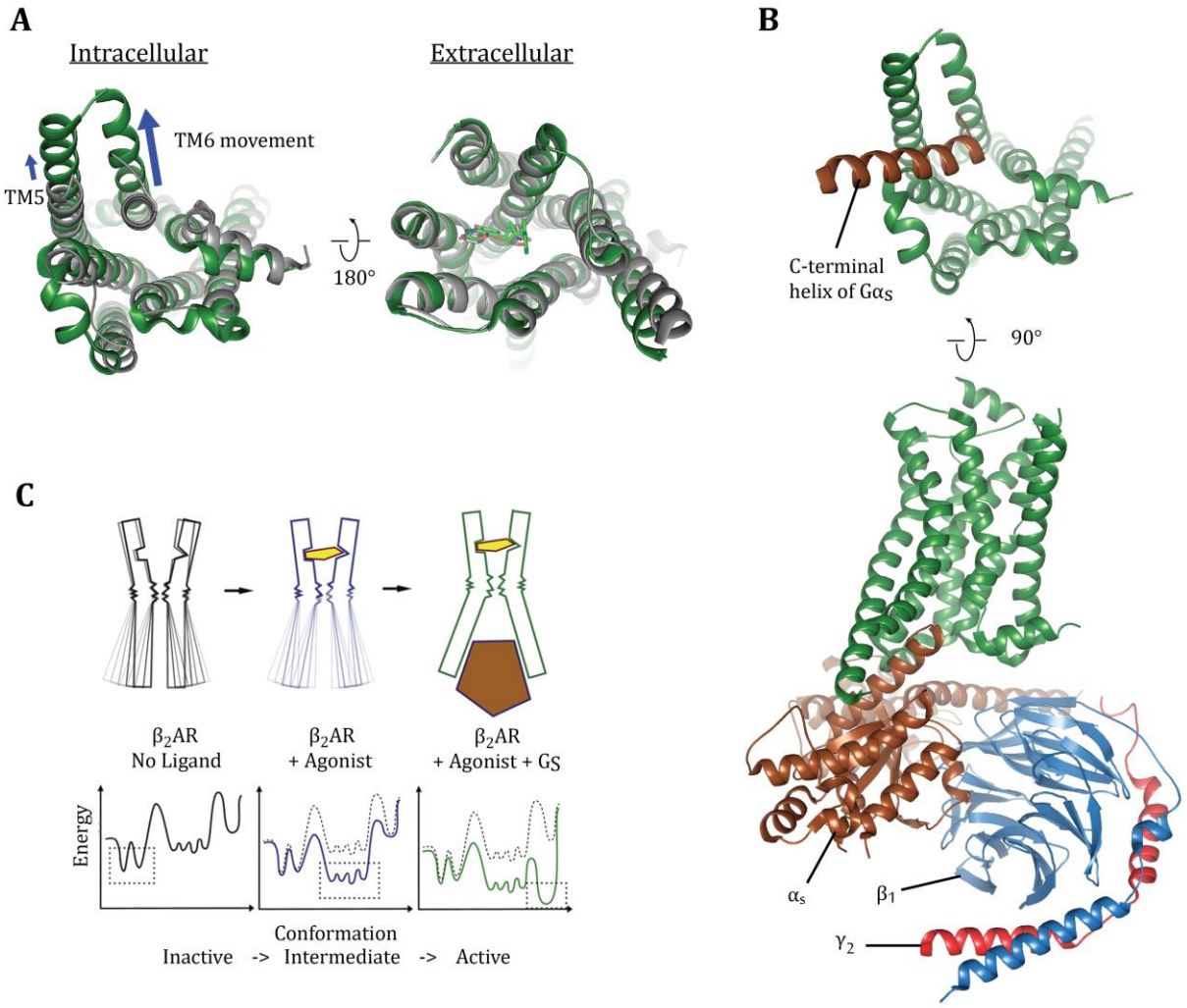


Figure 1.3 A, Current model of GPCR activation based on the example of β_2AR. A, Inactive (grey, carazolol-bound) and active (green, G_s- and irreversible full agonist-bound) receptor superposition (PDB-IDs: 2RH1, 3SN6). Note the large movement of TM5 and 6 at the cytosolic side (only observed in the presence of G_s) and the minor changes at the extracellular side. B, Receptor-$G\alpha_s\beta_1\gamma_2$ complex. For clarity, only the C-terminal helix of Gα is shown at the intracellular perspective (top). C, Suggested mechanism of β_2AR activation. The mobility of cytosolic regions increases and extracellular regions sample less conformations upon agonist binding. The G protein stabilizes the active state. Note that the suggested destabilizing effect of the agonist on the inactive state is not evident from the energy landscape scheme. Panel C is an adapted Illustration from Nygaard et al. 2013 [30].

1.2.1. Neurotensin receptor 1

Neurotensin is a tridecapeptide (pGlu-Leu-Tyr-Glu-Asn-Lys-Pro-Arg-Arg-Pro-Tyr-Ile-Leu) that was originally isolated from bovine hypothalamus in 1973[33] and subsequently from bovine intestine[34]. Like many neuropeptides, it fulfills a dual function as a neuromodulator in the central nervous system and as a local hormone in the periphery. In the brain, it exerts potent hypothermic and analgesic effects and it acts as an appetite suppressant. In addition, it modulates dopamine transmission and anterior pituitary hormone secretion[35]. It plays a role in

neuropsychiatric and neurodegenerative diseases[36]. In peripheral organs, neurotensin is an endocrine and paracrine modulator of the digestive tract and of the cardiovascular system. Furthermore, it was shown to act as a growth factor on cancer cells, suggesting that medical compounds targeting its signaling pathway are of high pharmacological relevance[35].

Three neurotensin receptors were identified (NTR1, NTR2, NTR3). NTR1 and NTR2 are class A GPCRs, whereas NTR3 (also termed sortilin[37]) is a single-transmembrane, multi-domain protein. Most of the described effects of neurotensin are mediated by NTR1, which was shown to interact with the six C-terminal amino acids of neurotensin (NT8-13). Northern blot analysis confirmed NTR1 expression in rat and human brain and intestine[35; 38]. Specifically in the central nervous system, high levels of NTR1 mRNA were found in several areas, including nucleus basalis magnocellularis, suprachiasmatic nucleus, substantia nigra and in the neurons of the diagonal band of Broca.

NTR1 interacts with several different signaling partners. It was shown to activate phospholipase C via coupling to G_q and to modulate adenylyl cyclase activity via G_i and G_s[40]. Moreover, in primary cultured rat brain neurons, in neuronal cell lines and in CHO and COS cells, it was demonstrated that NTR1 internalizes upon agonist binding[41]. Internalization was impaired by the double-mutation at Thr422 and Tyr424 in the C-terminal tail, suggesting that these residues may be relevant for GRK or β-arrestin interactions[42].

The broad spectrum of interacting signaling partners is typical for GPCRs and considering the availability of several synthetic ligands of alternative pharmacologies and the medical potential of NTR1 targeting strategies, it is clear that this receptor is a promising GPCR representative for in-depth studies including structural investigations.

1.3. Membrane protein structure determination

Understanding the molecular basis of life crucially depends on the availability of structural information at atomic resolution. In comparison to other biomolecules, high-resolution structural data on membrane proteins (MPs) is particularly scarce. Given that approximately 26 % of human genes are encoding MPs[43], it is striking that only 417 unique structures are available in the Protein Data Bank (September 2013)[44]. This contrasts to the total number of deposited, unique biomolecular structures, which is 20'879. Hence, it is evident that specific technical problems, related to MPs, are limiting structural characterizations.

The most commonly applied technique for structure determination at atomic resolution is x-ray crystallography[45]. MP-specific difficulties in crystallographic projects are low heterologous overexpression levels and reduced stability in detergent solubilized-states compared to their

native membrane environment. Both problems hamper preparation of milligram quantities of homogeneous and active protein samples and thus limit the number of crystallization conditions that can be screened. Additional problems are related to the application of detergents for MP handling. High amounts of detergents trigger disadvantageous phase separations in vapor diffusion crystallization experiments[46] and the presence of the detergent belt around transmembrane regions is a main reason, why MP crystals tend to have high solvent contents and poor diffraction[47].

Miniaturization and the development of high-throughput crystallization robotics has facilitated biomolecular crystallography in general ways in recent years. At the same time, MP-specific challenges were also tackled: First, by the development of novel membrane mimics, second, by the development of alternative crystallization methods and third, by the establishment of screening and protein engineering strategies that allow identification of suitable MPs for structural studies.

1.3.1. Alternative membrane mimics and the special case of GPCRs

The most commonly applied amphiphiles for extraction of MPs from membranes and for mimicking the native hydrophobic environment during purification are detergents. Newstead et al. provided in 2007 an extensive analysis of published MP crystallization conditions and reported, among other findings, that the majority of published structures were obtained using n-dodecyl-β-D-maltoside (DDM)[48]. The impression of DDM as a successful crystallization detergent may be biased to some extent, due to the fact that many MP crystallographers traditionally apply it as a first choice in crystallization trials[49; 50]. Nevertheless, it is apparent that DDM is one of the mildest available detergents and hence a promising candidate for crystallographic purposes.

In a more recent report, Chae et al. (2010) introduced a novel type of detergent, which is built around a central quaternary carbon atom derived from neopentyl glycol and two hydrophilic head groups derived from maltose[51]. These, so called maltose-neopentyl glycols (MNGs), tend to have very low critical micelle concentration (e.g. 0.001 % (w/v) for MNG-3) and were described to be efficient solubilization agents. Further, Chae et al. reported that MNGs can provide superior stability to various types of membrane proteins (relative to DDM), including GPCRs, a succinate:quinone oxidoreductase and a light-harvesting complex[51]. Based on the application of MNG-3 during purification, several crystal structures of GPCRs were published in the last three years[24; 52-59] and given the limited amount of time since Chae's initial publication, it is likely that MNGs will soon be established as an important tool in MP research.

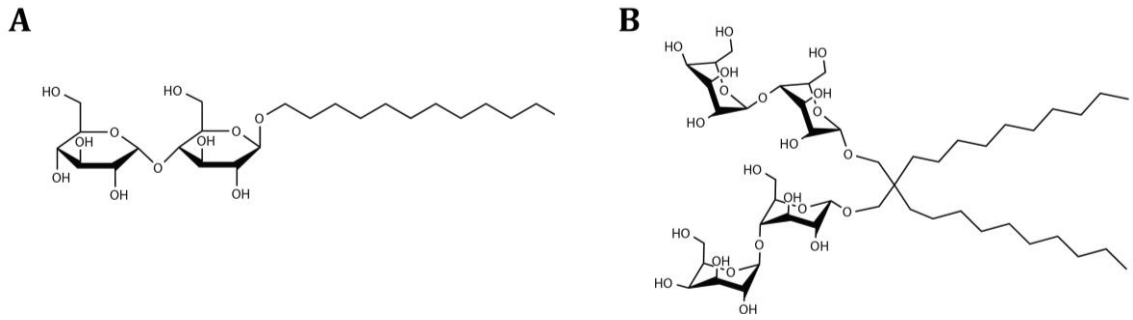


Figure 1.4 Chemical structures of n-dodecyl-β-D-maltopyranoside (A) and of lauryl neopentyl glycol (B).

In spite of the preference of MP crystallographers for DDM, this particular detergent has never been successfully used for the crystallization of GPCRs in vapor diffusion experiments. This was true at the very beginning of this thesis in 2009, when only four unique structures were available (rhodopsin, β_1-adrenergic receptor, β_2-adrenergic receptor, A2A adenosine receptor), and it is now still the case after the publication of more than 20 different receptor structures[60]. One explanation for the particular behavior of GPCRs in crystallization trials is related to the unusually small intra- and extracellular regions of most of these receptors. Since DDM forms large micelles, it is likely that the hydrophilic regions of GPCRs are partially covered by the detergent and thus occluded from crystal contact formation.

1.3.2. Lipidic cubic phase crystallization

One main strategy that was pursued to circumvent shielding effects of large micelles, was the application of lipidic cubic phase (LCP) crystallization, as it allows or favors crystal contact formation via transmembrane regions[61; 62]. Typically, reconstitution into the three-dimensional lipidic bilayer (the cubic phase) is performed using liquid monoolein and pure GPCRs in a mild detergent, such as MNG-3 or DDM. At 20°C, the cubic phase Pn3m forms spontaneously, when water (or aqueous protein solution) is mixed to liquid monoolein at the exact ratio of 1 to 1.5 [63]. Pn3m is bicontinuous, in the sense that it allows diffusion of solvents and of MPs in all 3 dimensions – both are crucial requirements for crystallogenesis[23]. Once LCP is formed, it can co-exist with excess aqueous solution (it does not dissolve), which is essential for crystal screening purposes. Hundreds of individual LCP boluses are usually covered by different crystallization solutions and the various solutes (salts, precipitants, buffering agents, etc.) diffuse into the LCP at different rates, thus exposing the reconstituted GPCRs to an array of changing solvent conditions over time. Interestingly, small-angle x-ray scattering experiments have clearly proven that the lipidic phase is indeed cubic after reconstitution and prior to crystallization. Nevertheless, the crystallized GPCRs reside always in a lamellar phase *locally*, which is evident from the exclusively parallel or antiparallel alignments of the GPCR transmembrane regions in

the crystal lattice (Figure 1.5). It is currently unclear, whether crystallization triggers the local transition to the lamellar phase or vice versa.

The alignment of the transmembrane segments and the crystal contacts via these regions are important reasons for the comparably small solvent content that LCP crystals often exhibit. And this may in turn explain why crystals of this type frequently diffract to relatively high resolution (e.g. the structure of A2A adenosine receptor at 1.8 Å at a solvent content of 51 %[64]). A clear disadvantage is, however, that it is challenging to obtain datasets with "high" completeness (> 90 % in the highest resolution bin). In many cases it was necessary to merge several small datasets, each covering only very few degrees of rotation. This has two LCP-specific reasons: First, LCP crystals are often significantly more sensitive to radiation damage compared to crystals from vapor diffusion experiments (personal communication, William I. Weis and Kaspar Hollenstein), which reduces the applicable x-ray dose. In my opinion, this might be due to higher diffusion rates of radicals within the membranous phase compared to the aqueous phase in vapor diffusion experiment (more secondary radiation damage). And second, the goniometric crystal centering in the x-ray beam is often very difficult, due to the uneven LCP surface in the loop and the different refractive indices of air and LCP (crystals are almost impossible to see after harvesting and mounting to the beam). The crystals have therefore often to be localized by x-ray test exposures using a dens raster, which causes additional radiation damage[65]. As a result of difficulties in goniometric centering, crystals can exit the x-ray beam after very few degrees of rotation during data collection.

The major disadvantage of the LCP methodology for GPCR crystallization is certainly the necessity of at least one soluble protein (or domain) to bridge the aqueous channels of the lipidic phase for crystal contact formation. In the vast majority of published structures, the intracellular loop 3 (IC3) was replaced by T4 lysozyme (T4L). This rendered the receptors signaling-inactive in all cases, as it prevented the interaction with the name-giving signaling partners (the G proteins). In one exceptional case, T4 lysozyme was attached to the GPCR N-terminus, but these crystals diffracted only to 4 Å, which might be due to the higher flexibility of the construct compared to the IC3 fusions[66]. The majority of high-resolution LCP structures are thus only obtained from GPCR-like hybrid molecules, not from *G protein-coupled* receptors and it is important to develop alternative strategies, which focus on the structure determination of signaling-active proteins. Structures of this type would likely shed light on novel aspects of GPCR signaling, due to a more reliable description of intracellular regions.

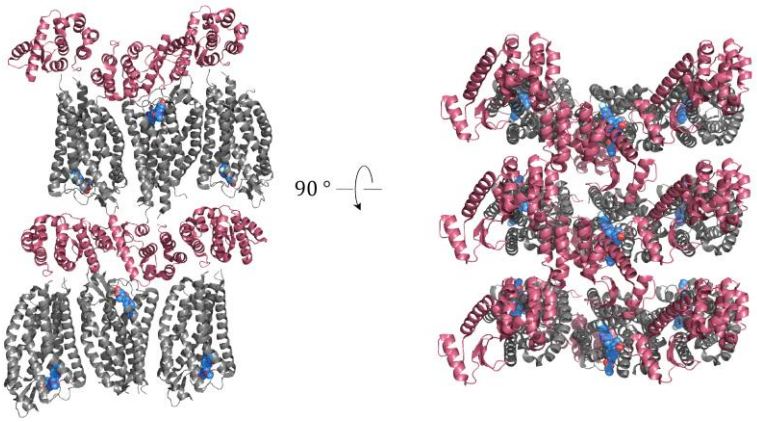


Figure 1.5 Example of LCP crystal packing (PDB-ID: 3VW7). Protease activated receptor 1 (PAR1, grey) fused to T4L (red) in complex with the antagonist vorapaxar (blue)[67]. These crystals diffracted to 2.2 Å and displayed a solvent content of 50.7 %, which is comparable to many soluble proteins (43 % on average[68]). The data exhibit a completeness of only 85 % in the highest resolution bin, in spite of the fact that datasets from 18 different crystals were merged. The space group is $P2_12_12_1$.

1.3.3. Screening the diversity of life and alanine-scanning

Advances in MP structural biology are not only due to the development of novel membrane mimics and because of improved LCP crystallization methods. Significant progress was also made at the level of the target proteins themselves in recent years. As mentioned above, low overexpression levels and reduced stability in detergent solution appear to be major factors that specifically counteract MP structure determination efforts. Both of these factors are protein-specific characteristics, and hence, sequence-dependent features.

A successful strategy for the identification of well-expressed and stable membrane proteins for structural biology is based on homology screening. Homology screens take advantage of the existing diversity of life, its common descent and the conservation of structure-function relationships at the molecular level during evolution. Given a specific human MP of interest, researchers use the available genomic data to extract *in silico* an array of homologs by BLAST algorithms. Efficient methodologies have been established that allow cloning of hundreds of homologous sequences from various organisms into expression vectors[69; 70], thus providing a basis for production and subsequent thermostability assessments or purification tests in high-throughput formats. Novel high-resolution structures of previously poorly characterized membrane protein families, such as ion channels and ABC transporters, could be determined in this way[7172; 73]. Even if a homolog under investigation is only distantly related to the protein of interest, the overall architecture can be similar and the conservation of key functional residues is often a valuable basis to conclude on certain structural aspects of human counterparts and on their molecular mechanisms[71; 74-77].

Homology screens depend on the existence of related genes that have evolved over the last billions of years and on the availability of their sequences. By far the most genomes that were sequenced in recent years are of prokaryotic origin (approximately twenty times more than eukaryotic[78]) and it is thus not surprising that mostly structures of prokaryotic homologs could be solved using this approach. Another explanation for this fact may be the higher variability of environmental conditions in prokaryotic niches compared to the controlled and thus constant settings of the intercellular spaces in multicellular eukaryotes. Some prokaryotic MPs may have been naturally selected to exhibit high stabilities in order to withstand harsh conditions, such as high salt concentrations (halophiles) or hot springs (thermophiles), which would be favorable for structural studies. Based on these considerations, it appears that the vast majority of eukaryotic MP families are not promising targets for homology screens, as no related prokaryotic MP genes were identified so far (only 256 common families were described out of a total of 1762 eukaryotic MP families[79]).

GPCRs, for example, which are uniquely eukaryotic MPs, may be very challenging targets for homology screening. But given their physiological importance, alternative strategies, such as alanine-scanning mutagenesis, were recently developed[80]. Using the alanine-scanning approach, the stability problem of GPCRs is tackled by the artificial introduction of hundreds of individual point mutations. Every wild-type residue of a GPCR (> 300 amino acids) is usually replaced by alanine (or leucine if alanine is the wild-type residue) and the resulting mutants are individually expressed and analyzed for thermostability in detergent solution. The most thermostable point mutations are subsequently combined in different pairs at the gene level and a novel round of expression and thermostability analysis is started. This process is repeatedly carried out until mutants with reasonably increased thermostability can be isolated, which are then subjected to structural studies.

The alanine-scanning methodology has lead to the structures of several GPCRs, mostly using the mild LCP crystallization methodology and lysozyme fusions[60; 80]. One drawback of this method is however, that – despite of considerable mutagenesis and screening efforts – it is often challenging to increase GPCR thermostabilities sufficiently for vapor diffusion crystallization. Instead, residual signaling-activity (if any after mutagenesis) is entirely abolished by the inactivating lysozyme, which critically promotes LCP crystallization. Since the crystallized constructs are frequently signaling-inactive, many structural aspects – mostly those distant to the ligand – are not reliably representing functionally relevant GPCRs states[19]. Further, alanine-scanning does not specifically tackle overexpression problems in heterologous systems, which is frequently an important limitation for GPCR structural research.

1.3.4. Directed evolution for structural studies of GPCRs

An interesting approach that focused on the GPCR overexpression problem was introduced by Casim Sarkar and other members of the Plückthun group in 2008[81]. *E. coli* – being the most frequently applied and simple heterologous expression host for proteins in general – had never been successfully used for GPCR structure determination at that time. It was apparent that its inapplicability for the functional production of large quantities of receptors hampered scientific progress in the GPCR field. The basic motivation of Sarkar et al. was therefore to develop a method that can generate altered wild-type GPCR amino acid sequences with higher functional expression levels in *E. coli*.

Briefly, genetic diversity was generated analogously to natural processes by random mutagenesis – albeit with a very high mutation rate, using error-prone PCR on a wild-type receptor gene. Approximately 10^9 different GPCR variants, carrying up to four different random mutations were subsequently transformed at once into *E. coli*. They were encoded on an expression vector, which provided an N-terminal maltose-binding protein (MBP) with a periplasmic signal sequence and an IPTG-inducible promotor. After induction of the cell pool and gentle permeabilization of the outer membrane, a fluorescent GPCR ligand was applied, which labelled preferentially those cells with high functional expression levels. This resulted in distinguishable phenotypes that were linked to their corresponding genotypes – a core requirement of evolution. As a selection pressure, fluorescence-activated cell sorting (FACS) was applied, which enriched the most expressing cells from a population of approximately 100 million "individuals" (10^8 cells/h can be sorted). This mutagenesis and selection cycle was repeatedly carried out and generated an array of GPCR variants with high functional expression levels in *E. coli* (Figure 1.6).

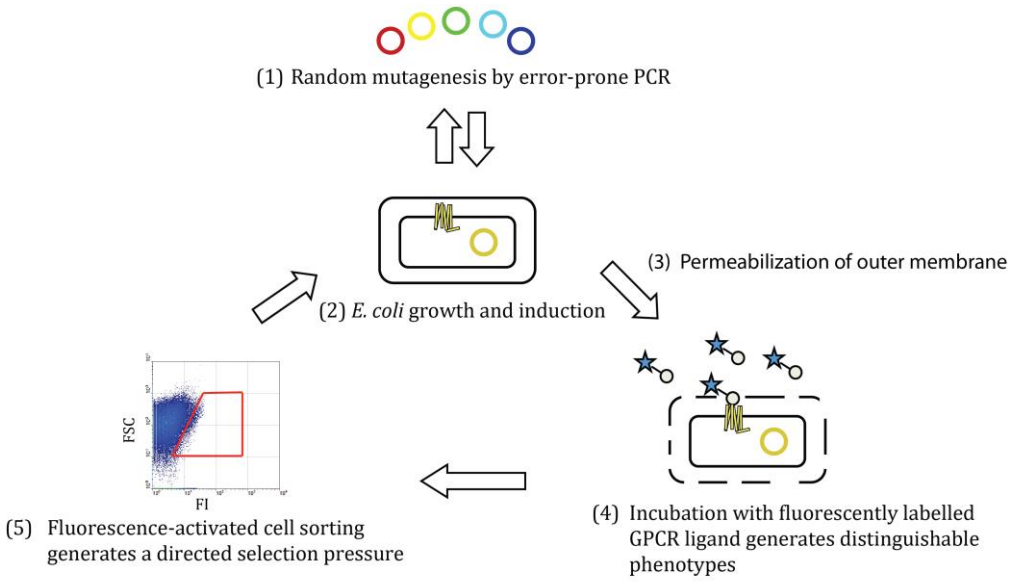


Figure 1.6 Directed evolution approach that improved the expression levels of neurotensin receptor 1 (10-fold), tachykinine receptor NK1 (>10-fold), α_{1A} adrenergic receptor (>10-fold) and α_{1B} adrenergic receptor (<10-fold). One *E. coli* cell of a library containing approximately 10^9 members is depicted at stage (2) and (4). Its inner and outer membranes are illustrated by 2 black smoothened rectangles. An expressing GPCR variant and its corresponding plasmid are depicted in yellow. The fluorescent ligand (black empty circle) is labeled with a fluorescent dye (blue asterisk). The FACS readout is forward scatter (FSC, a measure of cell size) and fluorescence intensity (FI). This figure was adapted from Sarkar et al. (2008)[81; 82]

Although a weak correlation between functional expression levels and higher thermostability was observed during directed evolution of certain GPCRs[82], extensive screening and mutant recombination efforts were still necessary to identify receptor variants with very high thermostabilities that were suitable for structure determination[83; 84]. Therefore, Daniel J. Scott from this laboratory adapted the above-described directed evolution method to specifically select for more stable receptors in detergent solution (Figure 1.7)[85]. The key development of this second technology, termed CHESS (Cellular High-througput Encapsulation Solubilization and Screening), was to surround *E. coli* cells after GPCR expression by a semipermeable polysaccharide capsule, which allowed solubilization of the cells by harsh detergents without loss of the genotype-phenotype linkage. Labelling of the capsules with a fluorescent ligand and the FACS-based selection pressure could subsequently be applied analogously to the method described above. Since the capsules cannot be cultivated like cells after enrichment, the genes encoding for the most stable receptor variants were amplified by PCR after FACS sorting. Subsequently, they were ligated to expression vectors before the next "generation" of directed evolution was started.

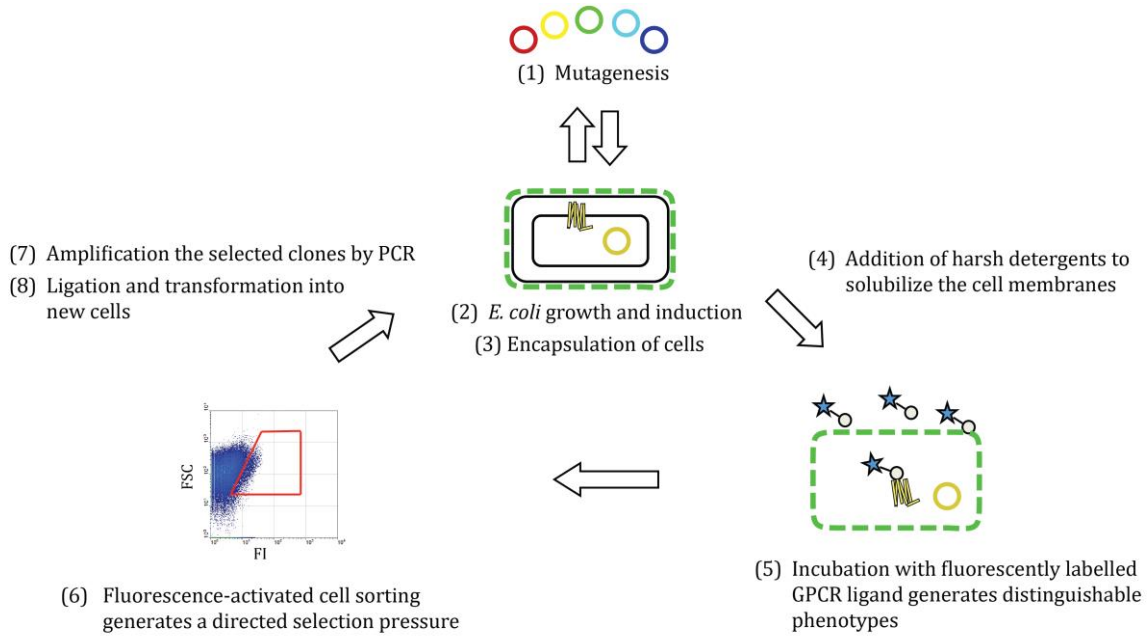


Figure 1.7 Directed evolution method for the generation of stable GPCR variants in harsh short-chain detergents. The detergent resistant capsule of *E. coli* is depicted in green. The encapsulation method takes advantage of the negative surface charge of *E. coli*, which aids in the deposition of positively charged chitosan (poly-D-glucosamine) and subsequently of a layer of negatively charged alginate. The pore size of the semipermeable capsule can be controlled by the number of deposited layers of alternative charges and it is fine-tuned to be permeable for detergent monomers and ligands and impermeable for solubilized receptors and plasmids. The figure was adapted from Sarkar et al. (2008)[81; 85].

Similar to the alanine-scanning approach, a limitation of directed evolution is the requirement for an MP ligand as a functional readout. Ligands have to be labelled fluorescently and they should be sufficiently small to diffuse into the periplasmic space or through the pores of the capsules. They have to be soluble in water and should exhibit little non-specific binding to cellular components or to the capsules. These requirements limit the current applicability of directed evolution to receptors with suitable ligands, suggesting that further developments are needed to increase the target range of this strategy and to make it generally applicable for all MPs.

In spite of this limitation, it is apparent that directed evolution bears an enormous potential for the efficient generation of useful MP variants, since it does not rely on manual trial-and-error procedures. Unlike homology screens or alanine-scans, this concept requires no handling of single MP variants – neither at the level of molecular cloning, nor at the phenotypic

characterization stage – as it relies on random mutagenesis and on the fully automated FACS procedure with two orders of magnitude higher throughput than 384-well plates.

Compared to alanine-scanning, which is restricted to the substitution of one type of amino acid, directed evolution benefits from the possibility to introduce the entire range of natural amino acids during random mutagenesis. This may explain why significantly higher thermostabilities of evolved variants were measured than for the best alanine mutants in the case of NTR1 (the only GPCR, where both methods were applied)[86].

Compared to homology screens, an advantage of directed evolution is that it does not require a pre-existing diversity of life, which was generated by variable selections and genetic drift over billions of years. It relies on a controlled selection pressure that can direct the evolution of specific characteristics within weeks. As a result, improved MP variants can be generated while retaining a very high sequence identity to the protein of interest. Thus, in contrast to most crystallized prokaryotic homologs of human proteins, evolved MPs can in principle be used to obtain structural data of direct pharmacological relevance.

1.4. Goal of this work

Directed evolution for high functional expression levels in *E. coli* was published by Sarkar et al.[81] before the start of my dissertation in June 2009. In this study, NTR1 – a GPCR that had never been crystallized in spite of decades of crystallographic efforts[87-92] – was used as a proof of principle case to demonstrate the feasibility of the evolutionary method. Sarkar et al. obtained an NTR1 variant, termed D03, which exhibited approximately 10-fold increased functional expression levels in *E. coli* and more than 98 % sequence identity to the wild-type GPCR of interest.

My aim was to solve a high-resolution crystal structure of this evolved NTR1 variant in complex with its natural agonist neurotensin. This had several reasons: First, a structure would prove that directed evolution can be useful for GPCR structural biology and potentially also for other MP targets. Second, *E. coli* would be introduced or established as a valuable host for functional GPCR overexpression at large quantities, which would pave the way for various NMR experiments tackling receptor dynamics. This had previously not been feasible for perdeuterated and functionally produced GPCRs from *E. coli*. And third, in 2009 only four unique GPCR structures were available – mostly showing receptors in inactive conformations. NTR1 bound to an agonist (neurotensin) would give novel insights about GPCR activation and since no peptide-binding GPCR structures were available, a structure of this type would shed light on new aspects of an entire subfamily of class A GPCRs.

1.5. Initial progress

1.5.1. Previous work

The group of Reinhard Grisshammer at the NIH membrane protein structure and function unit in Rockville (MD, USA) attempted to solve the structure of NTR1 since 1993 and published expression trials in *E. coli* and several improvements of a purification procedure[49; 50; 51; 52; 53; 54]. Their NTR1 isolation method relies on extraction from whole cells using a mixture of DDM/CHAPS/CHS, on immobilized metal-chelate affinity chromatography (IMAC) and subsequently, on a ligand-column that is based on the interaction of streptavidin and biotinylated neurotensin. A postdoctoral fellow and a PhD student tried to implement these purification principles in our group and started crystallization trials using the evolved variant D03. In spite of my extensive analysis of their laboratory documentations (they left the group before the start of my work), I was not able to find clear evidence for successful large-scale expression experiments or for reasonably pure GPCR samples and monodisperse size-exclusion chromatography (SEC) runs.

1.5.2. Implementation of the "Grisshammer protocol"

I started off by implementing and improving the efficiency of the individual steps of the Grisshammer protocol systematically one by one for the D03 variant. My initial experiments focused on expression optimization using different cultivation temperatures, *E. coli* strains and induction conditions. *E. coli* BL21 Tuner cells cultivated at 29 °C showed higher functional expression levels per cell in radioligand-binding assays compared to *E. coli* DH5α cultivated at 20 °C (Grisshammer protocol). In addition, *E. coli* BL21 Tuner cells grew reproducibly to significantly higher cell densities under expression conditions. This lead to more than 2-fold increased functional expression using this alternative strain, which corresponded to a total of approximately 1.6 mg of active D03 per liter of culture, as estimated by radioligand-binding measurements.

Subsequently, western blots, radioligand-binding assays and SDS-PAGE analyses were used to monitor solubilization and purification efficiencies (a selection of results are shown in Figure 1.8). The only adaption of the Grisshammer protocol in these trials was the replacement of the streptavidin matrix by a monomeric avidin resin. This was necessary due to the reduced salt sensitivity of D03 in terms of ligand-binding. In the Grisshammer protocol, NTR1-wildtype is eluted from the biotinylated neurotensin (immobilized at the streptavidin resin) by high salt concentrations. This was not possible for D03 even in the presence of excess NT8-13 (Figure 1.8C), as ligand-binding appeared to be salt-insensitive. The monomeric avidin resin, which we

applied, was shown to display a weaker interaction to biotin than streptavidin, so that biotinylated ligands can be eluted using excess biotin[93]. In contrast to the elution procedure in the Grisshammer protocol, which disrupts the GPCR-ligand interaction, we could thus attempt to elute the receptor-ligand complex as a whole.

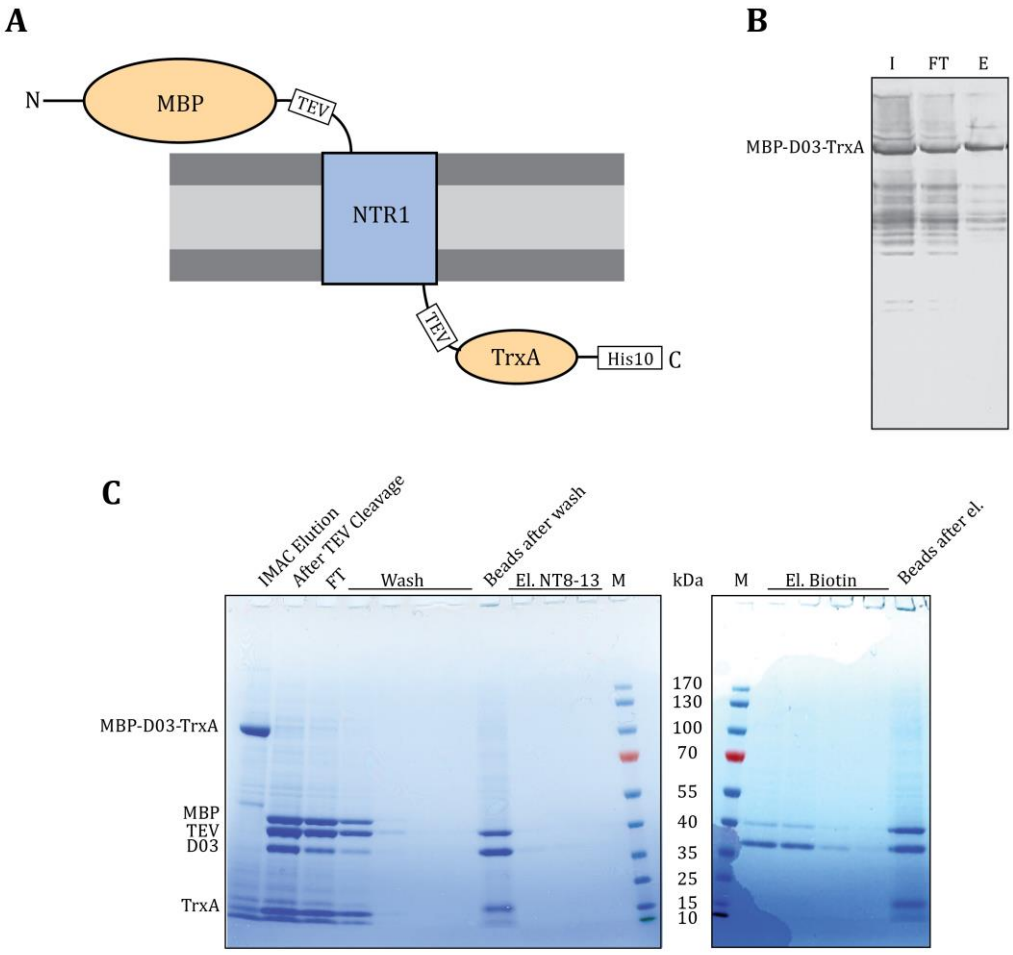


Figure 1.8 D03 purification trials according to the protocol of the Grisshammer group (previously developed for wild-type NTR1). A, *E. coli* expression construct. B, Western blot analysis of a representative IMAC purification using an anti-MBP primary antibody. C, SDS-PAGE of key fractions during monomeric avidin/biotinylated neurotensin purification (progression from left to right). Note that high salt elution using excess neurotensin (El. NT8-13) was not successful, but immobilized D03 could be eluted by competition with excess biotin from the monomeric avidin resin (El. Biotin). A major fraction of D03, however, remained on the resin, suggesting that it bound non-specifically.

The major problems of the purification trials according to the Grisshammer protocol were the time requirements until completion of the procedure (4 days) and the low yields after the final ligand-affinity step (< 0.08 mg per liter *E. coli* culture). The latter was due to the following three reasons: First, in spite of extensive IMAC tests regarding incubation time during binding, imidazole concentrations and column materials, it was not possible to isolate a major fraction of D03 – a large part of the fusion protein was in the IMAC flow-through (Figure 1.8B is a

representative example). The column capacity of the Ni-NTA IMAC resin (QUIAGEN) for the D03 fusion construct was estimated to be below 0.1 mg/ml. Second, TEV protease cleavage was highly inefficient, as stoichiometric amounts and overnight incubation were necessary to achieve complete cleavage (Figure 1.8C). And third, only a small fraction of D03 could be eluted from the monomeric avidin beads, independent of the applied biotin concentration or incubation time (biotin was usually applied in large excess). The major fraction remained bound to the monomeric avidin resin (Figure 1.8C), implying that D03 interacted mostly in a non-specific way with the column material, which might have been due to its aggregation tendency.

1.5.3. An alternative purification procedure

Inspired by a generally applicable membrane protein isolation procedure that is frequently carried out in the Group of Raimund Dutzler (Institute of Biochemistry University of Zurich), it was also attempted to purify D03 from prepared membranes using IMAC, followed by "reverse IMAC" after protease cleavage and imidazole removal. The main problems using this approach were again a relatively low binding efficiency during the initial IMAC step, the inefficient TEV protease cleavage reaction (typically performed with rhinovirus 3C protease in the Dutzler group) and loss of most of the D03 in the reverse IMAC step due to column interaction (in spite of complete His-tag removal, data not shown). One attempt was made to replace the reverse IMAC step by cation exchange chromatography, which resulted in smaller losses of D03. However, due to the inefficient binding to the IMAC in the first purification step of the protocol, the final yields were still low and the purified proteins appeared to be mostly aggregated, as judged by SEC (Figure 1.9).

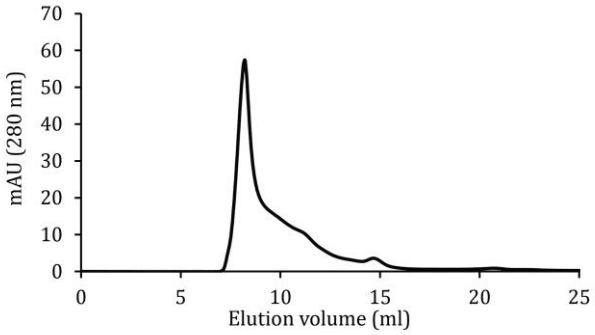


Figure 1.9 SEC analysis of D03 in 0.1 % (w/v) DDM. This chromatogram was obtained after extraction from prepared membranes, followed by the subsequent steps IMAC, TEV protease cleavage and cation exchange. The void volume of this column (Superdex 200 10/300 GL) is at an elution volume of 8 ml. Monomeric D03 in a DDM micelle would be expected to elute between 12 – 14 ml. SDS-PAGE analysis of collected fractions implied that the void peak consists mainly of D03 and to a lesser extent of residual fusion proteins (cleaved) and of TEV protease (data not shown).

1.5.1. Conclusions and further strategy development

The various observations of non-specific binding events, the surprisingly low capacity of IMAC resins and the poor running behavior on S200 were consistent with D03 being an unstable protein. Radioligand-binding measurements of purified D03 in DDM further indicated a half-life of less than 10 h at 4 °C, which was clearly too short for crystallographic purposes. As the ligand column was the last step of the protocol and since it took place more than 48 hours after solubilization, this finding explained well, why only a small fraction of the functionally expressed receptor could be eluted as a ligand-bound complex from the monomeric avidin resin.

The previously obtained structures of rhodopsin, β_1-adrenergic receptor, β_2-adrenergic receptor, A2A adenosine receptor were all determined in ligand-bound states, which likely reflected the fact that ligand binding frequently increases the stability of its binding protein. Considering this, it was investigated whether ligand-binding could potentially prevent the fast decay of D03 activity in solution, which was indeed the case (Figure 1.10)

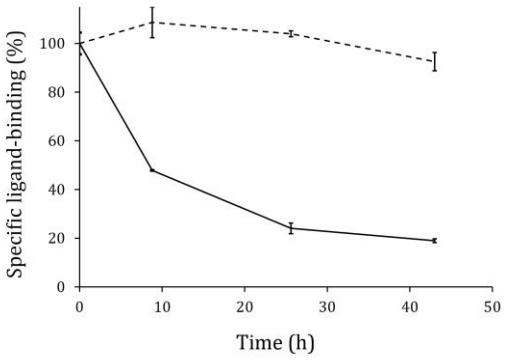


Figure 1.10 Stability measurement of purified protein in solution (10 mM HEPES pH8, 200 mM NaCl, 300 mM imidazole, 0.1 % (w/v) DDM). 1 ml SEC columns were used to separate bound radioligand (void) from free (retained within the column) in a table-top centrifuge. The solid line represents specific ligand-binding activity after incubation in the absence of ^{3}H-neurotensin, whereas the dashed line was obtained after incubation periods in its presence.

In summary, my initial experiments generated a body of evidence that consolidated previous observations on D03[81], but importantly, they also outlined new core problems for structure determination on evolved NTR1 variants and an improvement in this regard: First, it was clear that functional expression levels (1.6 mg/L of *E. coli* culture) were no longer limiting crystallographic structure determination attempts. Second, unlike previously stated[81], the fast decay of ligand-binding activity in solution suggested that D03 may not be of sufficient stability for crystallization. Third, the stability in DDM could be improved significantly by ligand binding. And fourth, since the ligand-affinity column is implemented as a *last* step in the Grisshammer

protocol and because of the long processing time, the procedure appeared to be suboptimal for crystallization.

An extensive literature research further showed that significantly harsher detergents than DDM are needed to be applied for GPCRs, if structure determination in the absence of T4 lysozyme (by vapor diffusion) is attempted (e.g. OG, NG or HEGA10). Based on the sum of these findings, a new tripartite strategy was followed:

1. I focused on the development of a more efficient purification strategy, which is based on a ligand column as the initial isolation step. A procedure of this type would profit from an early stabilization effect of the ligand, but at the same time, it would still allow separation of native and unfolded receptors. In addition, I generated various NTR1 constructs with alternative truncations for crystallographic purposes.

2. Karola Schlinkmann from our group applied directed evolution for functional expression and subsequently, thermostability screening and recombination procedures to identify mutants with higher thermostability[84; 86]. The variant TM86V, which was crystallized in the course of my work was obtained as a result of her progress.

3. Daniel James Scott from our group used the opportunity of lacking stability to develop and apply CHESS as described in section 1.3.4. [85]. The variants OGG7 and HTGH4 that were crystallized in the course of my work were obtained as a result of his progress.

1.6. Thesis outline

My own work included functional assays on NTR1 mutants, but mostly purification optimization, crystallization and structure determination of evolved variants. Chapter 2 of this thesis is a manuscript focusing on a novel purification strategy in the GPCR field that was developed in the course of this work (to be submitted). The described purification strategy was the basis for all successful NTR1 structure determination attempts in our research group. Chapter 3 is a PNAS article that summarizes our structural data on the 3 crystallized NTR1 variants and its implications for our understanding GPCR function. Chapter 4 is meant to be a brief outlook towards studying dynamic aspects of GPCR signaling by means of our well-behaved model proteins. It comprises initial results from solution-state NMR, which were obtained in collaboration with Dr. Franz Hagn (Wagner group, Harvard Medical School).

1.7. References

1. Mora, C., Tittensor, D. P., Adl, S., Simpson, A. G. & Worm, B. (2011). How many species are there on Earth and in the ocean? *PLoS Biol.* **9**, e1001127.
2. Darwin, R. C. (1859). *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life.* (Murray, J., Ed.), London.
3. Mendoza A, S.-P. A., Ruiz-Trillo I. (2014). The evolution of the GPCR signalling system in eukaryotes: modularity, conservation and the transition to metazoan multicellularity. *Genome Biol. Evol.* **6**, 606-619.
4. Kroeze, W. K., Sheffler, D. J. & Roth, B. L. (2003). G-protein-coupled receptors at a glance. *J. Cell Sci.* **116**, 4867-4869.
5. Fredriksson, R., Lagerstrom, M. C., Lundin, L. G. & Schioth, H. B. (2003). The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. *Mol. Pharmacol.* **63**, 1256-1272.
6. Palczewski, K., Kumasaka, T., Hori, T., Behnke, C. A., Motoshima, H., Fox, B. A., Le Trong, I., Teller, D. C., Okada, T., Stenkamp, R. E., Yamamoto, M. & Miyano, M. (2000). Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* **289**, 739-745.
7. Dobner, P. R. (2005). Multitasking with neurotensin in the central nervous system. *Cell. Mol. Life Sci.* **62**, 1946-1963.
8. Shenoy, S. K. & Lefkowitz, R. J. (2005). Seven-transmembrane receptor signaling through beta-arrestin. *Sci. STKE* **2005**, cm10.
9. Wilkie, T. M., Gilbert, D. J., Olsen, A. S., Chen, X. N., Amatruda, T. T., Korenberg, J. R., Trask, B. J., Dejong, P., Reed, R. R., Simon, M. I., Jenkins, N. A. & Copeland, N. G. (1992). Evolution of the Mammalian G-Protein Alpha-Subunit Multigene Family. *Nat. Genet.* **1**, 85-91.
10. Hurowitz, E. H., Melnyk, J. M., Chen, Y., Kouros-Mehr, H., Simon, M. I. & Shizuya, H. (2000). Genomic characterization of the human heterotrimeric G protein alpha, beta, and gamma subunit genes. *DNA Res.* **7**, 111-120.
11. Morishita, R., Nakayama, H., Isobe, T., Matsuda, T., Hashimoto, Y., Okano, T., Fukada, Y., Mizuno, K., Ohno, S., Kozawa, O., Kato, K. & Asano, T. (1995). Primary Structure of a Gamma-Subunit of G-Protein, Gamma(12), and Its Phosphorylation by Protein-Kinase-C. *J. Biol. Chem.* **270**, 29469-29475.
12. Ray, K., Kunsch, C., Bonner, L. M. & Robishaw, J. D. (1995). Isolation of Cdna Clones Encoding 8 Different Human G-Protein Gamma-Subunits, Including 3 Novel Forms Designated the Gamma(4), Gamma(10), and Gamma(11) Subunits. *J. Biol. Chem.* **270**, 21765-21771.
13. Kitabgi, P. (2002). Targeting neurotensin receptors with agonists and antagonists for therapeutic purposes. *Current Opinion in Drug Discovery & Development* **5**, 764-776.
14. Yamada, M., Yamada, M., Watson, M. A. & Richelson, E. (1994). Deletion Mutation in the Putative 3rd Intracellular Loop of the Rat Neurotensin Receptor Abolishes Polyphosphoinositide Hydrolysis but Not Cyclic-Amp Formation in Cho-K1 Cells. *Mol. Pharmacol.* **46**, 470-476.
15. Bozou, J. C., Amar, S., Vincent, J. P. & Kitabgi, P. (1986). Neurotensin-Mediated Inhibition of Cyclic-Amp Formation in Neuroblastoma N1e115-Cells - Involvement of the Inhibitory Gtp-Binding Component of Adenylate-Cyclase. *Mol. Pharmacol.* **29**, 489-496.
16. Ishizuka, J., Townsend, C. M., Thompson, J. C., Hanks, J. B. & Robinson, D. S. (1993). Neurotensin Regulates Growth of Human Pancreatic-Cancer. *Ann. Surg.* **217**, 439-446.
17. Nordstrom, K. J. V., Almen, M. S., Edstam, M. M., Fredriksson, R. & Schioth, H. B. (2011). Independent HHsearch, Needleman-Wunsch-Based, and Motif Analyses Reveal the Overall Hierarchy for Most of the G Protein-Coupled Receptor Families. *Mol. Biol. Evol.* **28**, 2471-2480.
18. de Mendoza A, S.-P. A., Ruiz-Trillo I. (2014). The evolution of the GPCR signaling system in eukaryotes: modularity, conservation, and the transition to metazoan multicellularity. *Genome Biol. Evol.* **6**, 606-619.
19. Egloff, P., Hillenbrand, M., Klenk, C., Batyuk, A., Heine, P., Balada, S., Schlinkmann, K. M., Scott, D. J., Schutz, M. & Pluckthun, A. (2014). Structure of signaling-competent

neurotensin receptor 1 obtained by directed evolution in Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.* **111**, E655-E662.
20. Savdie, C., Ferguson, S. S. G., Vincent, J. P., Beaudet, A. & Stroh, T. (2006). Cell-type-specific pathways of neurotensin endocytosis. *Cell Tissue Res.* **324**, 69-85.
21. McDonald, P. H., Chow, C. W., Miller, W. E., Laporte, S. A., Field, M. E., Lin, F. T., Davis, R. J. & Lefkowitz, R. J. (2000). Beta-arrestin 2: a receptor-regulated MAPK scaffold for the activation of JNK3. *Science* **290**, 1574-7.
22. Ge, L., Ly, Y., Hollenberg, M. & DeFea, K. (2003). A beta-arrestin-dependent scaffold is associated with prolonged MAPK activation in pseudopodia during protease-activated receptor-2-induced chemotaxis. *J. Biol. Chem.* **278**, 34418-26.
23. Wisler JW, X. K., Thomsen AR, Lefkowitz RJ. (2014). Recent developments in biased agonism. *Curr. Opin. Cell Biol.*, 18-24.
24. Rasmussen, S. G. F., DeVree, B. T., Zou, Y. Z., Kruse, A. C., Chung, K. Y., Kobilka, T. S., Thian, F. S., Chae, P. S., Pardon, E., Calinski, D., Mathiesen, J. M., Shah, S. T. A., Lyons, J. A., Caffrey, M., Gellman, S. H., Steyaert, J., Skiniotis, G., Weis, W. I., Sunahara, R. K. & Kobilka, B. K. (2011). Crystal structure of the beta(2) adrenergic receptor-Gs protein complex. *Nature* **477**, 549-U311.
25. Choe, H. W., Kim, Y. J., Park, J. H., Morizumi, T., Pai, E. F., Krauss, N., Hofmann, K. P., Scheerer, P. & Ernst, O. P. (2011). Crystal structure of metarhodopsin II. *Nature* **471**, 651-U137.
26. Singhal, A., Ostermaier, M. K., Vishnivetskiy, S. A., Panneels, V., Homan, K. T., Tesmer, J. J. G., Veprintsev, D., Deupi, X., Gurevich, V. V., Schertler, G. F. X. & Standfuss, J. (2013). Insights into congenital stationary night blindness based on the structure of G90D rhodopsin. *Embo Reports* **14**, 520-526.
27. Standfuss, J., Edwards, P. C., D'Antona, A., Fransen, M., Xie, G. F., Oprian, D. D. & Schertler, G. F. X. (2011). The structural basis of agonist-induced activation in constitutively active rhodopsin. *Nature* **471**, 656-660.
28. Scheerer, P., Park, J. H., Hildebrand, P. W., Kim, Y. J., Krauss, N., Choe, H. W., Hofmann, K. P. & Ernst, O. P. (2008). Crystal structure of opsin in its G-protein-interacting conformation. *Nature* **455**, 497-U30.
29. Deupi, X. & Kobilka, B. K. (2010). Energy Landscapes as a Tool to Integrate GPCR Structure, Dynamics, and Function. *Physiology* **25**, 293-303.
30. Nygaard, R., Zou, Y. Z., Dror, R. O., Mildorf, T. J., Arlow, D. H., Manglik, A., Pan, A. C., Liu, C. W., Fung, J. J., Bokoch, M. P., Thian, F. S., Kobilka, T. S., Shaw, D. E., Mueller, L., Prosser, R. S. & Kobilka, B. K. (2013). The Dynamic Process of beta(2)-Adrenergic Receptor Activation. *Cell* **152**, 532-542.
31. Manglik A, K. B. (2014). The role of protein dynamics in GPCR function: insights from the beta 2 adrenergic receptor and rhodopsin. *Curr. Opin. Cell Biol.* **27**, 136 - 143.
32. Altenbach, C., Kusnetzow, A. K., Ernst, O. P., Hofmann, K. P. & Hubbell, W. L. (2008). High-resolution distance mapping in rhodopsin reveals the pattern of helix movement due to activation. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 7439-7444.
33. Carraway, R. & Leeman, S. E. (1973). The isolation of a new hypotensive peptide, neurotensin, from bovine hypothalami. *J. Biol. Chem.* **248**, 6854-61.
34. Kitabgi, P., Carraway, R. & Leeman, S. E. (1976). Isolation of a tridecapeptide from bovine intestinal tissue and its partial characterization as neurotensin. *J. Biol. Chem.* **251**, 7053-8.
35. Vincent, J. P., Mazella, J. & Kitabgi, P. (1999). Neurotensin and neurotensin receptors. *Trends Pharmacol. Sci.* **20**, 302-9.
36. McMahon, B. M., Boules, M., Warrington, L. & Richelson, E. (2002). Neurotensin analogs indications for use as potential antipsychotic compounds. *Life Sci.* **70**, 1101-19.
37. Sarret, P., Krzywkowski, P., Segal, L., Nielsen, M. S., Petersen, C. M., Mazella, J., Stroh, T. & Beaudet, A. (2003). Distribution of NTS3 receptor/sortilin mRNA and protein in the rat central nervous system. *J. Comp. Neurol.* **461**, 483-505.
38. Tanaka, K., Masu, M. & Nakanishi, S. (1990). Structure and functional expression of the cloned rat neurotensin receptor. *Neuron* **4**, 847-54.

39. Vita, N., Laurent, P., Lefort, S., Chalon, P., Dumont, X., Kaghad, M., Gully, D., Le Fur, G., Ferrara, P. & Caput, D. (1993). Cloning and expression of a complementary DNA encoding a high affinity human neurotensin receptor. *FEBS Lett.* **317**, 139-42.
40. Kitabgi, P. (2006). Functional domains of the subtype 1 neurotensin receptor (NTS1). *Peptides* **27**, 2461-8.
41. Beaudet, A., Mazella, J., Nouel, D., Chabry, J., Castel, M. N., Laduron, P., Kitabgi, P. & Faure, M. P. (1994). Internalization and intracellular mobilization of neurotensin in neuronal cells. *Biochem. Pharmacol.* **47**, 43-52.
42. Chabry, J., Botto, J. M., Nouel, D., Beaudet, A., Vincent, J. P. & Mazella, J. (1995). Thr-422 and Tyr-424 residues in the carboxyl terminus are critical for the internalization of the rat neurotensin receptor. *J. Biol. Chem.* **270**, 2439-42.
43. Fagerberg, L., Jonasson, K., von Heijne, G., Uhlen, M. & Berglund, L. (2010). Prediction of the human membrane proteome. *Proteomics* **10**, 1141-9.
44. Bernstein, F. C., Koetzle, T. F., Williams, G. J., Meyer, E. F., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977). The Protein Data Bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.* **112**, 535-42.
45. Bernstein, F. C., Koetzle, T. F., Williams, G. J., Meyer, E. F., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977). The Protein Data Bank. A computer-based archival file for macromolecular structures. *Eur. J. Biochem.* **80**, 319-24.
46. Wiener, M. C. (2004). A pedestrian guide to membrane protein crystallization. *Methods* **34**, 364-72.
47. Kantardjieff, K. A. & Rupp, B. (2003). Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. *Protein Sci.* **12**, 1865-71.
48. Newstead, S., Ferrandon, S. & Iwata, S. (2008). Rationalizing alpha-helical membrane protein crystallization. *Protein Sci.* **17**, 466-72.
49. Sonoda, Y., Newstead, S., Hu, N. J., Alguel, Y., Nji, E., Beis, K., Yashiro, S., Lee, C., Leung, J., Cameron, A. D., Byrne, B., Iwata, S. & Drew, D. (2011). Benchmarking membrane protein detergent stability for improving throughput of high-resolution X-ray structures. *Structure* **19**, 17-25.
50. Kang, H. J., Lee, C. & Drew, D. (2013). Breaking the barriers in membrane protein crystallography. *Int. J. Biochem. Cell Biol.* **45**, 636-44.
51. Chae, P. S., Rasmussen, S. G., Rana, R. R., Gotfryd, K., Chandra, R., Goren, M. A., Kruse, A. C., Nurva, S., Loland, C. J., Pierre, Y., Drew, D., Popot, J. L., Picot, D., Fox, B. G., Guan, L., Gether, U., Byrne, B., Kobilka, B. & Gellman, S. H. (2010). Maltose-neopentyl glycol (MNG) amphiphiles for solubilization, stabilization and crystallization of membrane proteins. *Nat. Methods* **7**, 1003-8.
52. Granier, S., Manglik, A., Kruse, A. C., Kobilka, T. S., Thian, F. S., Weis, W. I. & Kobilka, B. K. (2012). Structure of the delta-opioid receptor bound to naltrindole. *Nature* **485**, 400-U171.
53. Rasmussen, S. G. F., Choi, H. J., Fung, J. J., Pardon, E., Casarosa, P., Chae, P. S., DeVree, B. T., Rosenbaum, D. M., Thian, F. S., Kobilka, T. S., Schnapp, A., Konetzki, I., Sunahara, R. K., Gellman, S. H., Pautsch, A., Steyaert, J., Weis, W. I. & Kobilka, B. K. (2011). Structure of a nanobody-stabilized active state of the beta(2) adrenoceptor. *Nature* **469**, 175-180.
54. Kruse, A. C., Hu, J. X., Pan, A. C., Arlow, D. H., Rosenbaum, D. M., Rosemond, E., Green, H. F., Liu, T., Chae, P. S., Dror, R. O., Shaw, D. E., Weis, W. I., Wess, J. & Kobilka, B. K. (2012). Structure and dynamics of the M3 muscarinic acetylcholine receptor. *Nature* **482**, 552-556.
55. Manglik, A., Kruse, A. C., Kobilka, T. S., Thian, F. S., Mathiesen, J. M., Sunahara, R. K., Pardo, L., Weis, W. I., Kobilka, B. K. & Granier, S. (2012). Crystal structure of the mu-opioid receptor bound to a morphinan antagonist. *Nature* **485**, 321-U170.
56. Kruse, A. C., Ring, A. M., Manglik, A., Hu, J. X., Hu, K., Eitel, K., Hubner, H., Pardon, E., Valant, C., Sexton, P. M., Christopoulos, A., Felder, C. C., Gmeiner, P., Steyaert, J., Weis, W. I., Garcia, K. C., Wess, J. & Kobilka, B. K. (2013). Activation and allosteric modulation of a muscarinic acetylcholine receptor. *Nature* **504**, 101-+.

57. Haga, K., Kruse, A. C., Asada, H., Yurugi-Kobayashi, T., Shiroishi, M., Zhang, C., Weis, W. I., Okada, T., Kobilka, B. K., Haga, T. & Kobayashi, T. (2012). Structure of the human M2 muscarinic acetylcholine receptor bound to an antagonist. *Nature* **482**, 547-U147.
58. Rosenbaum, D. M., Zhang, C., Lyons, J. A., Holl, R., Aragao, D., Arlow, D. H., Rasmussen, S. G. F., Choi, H. J., DeVree, B. T., Sunahara, R. K., Chae, P. S., Gellman, S. H., Dror, R. O., Shaw, D. E., Weis, W. I., Caffrey, M., Gmeiner, P. & Kobilka, B. K. (2011). Structure and function of an irreversible agonist-beta(2) adrenoceptor complex. *Nature* **469**, 236-240.
59. Ring, A. M., Manglik, A., Kruse, A. C., Enos, M. D., Weis, W. I., Garcia, K. C. & Kobilka, B. K. (2013). Adrenaline-activated structure of beta(2)-adrenoceptor stabilized by an engineered nanobody. *Nature* **502**, 575.
60. Katritch, V., Cherezov, V. & Stevens, R. C. (2013). Structure-Function of the G Protein-Coupled Receptor Superfamily. *Annu. Rev. Pharmacol. Toxicol.* **53**, 531-556.
61. Cherezov, V., Abola, E. & Stevens, R. C. (2010). Recent Progress in the Structure Determination of GPCRs, a Membrane Protein Family with High Potential as Pharmaceutical Targets. *Methods Mol. Biol.* **654**, 141-168.
62. Xu, F., Liu, W., Hanson, M. A., Stevens, R. C. & Cherezov, V. (2011). Development of an Automated High Throughput LCP-FRAP Assay to Guide Membrane Protein Crystallization in Lipid Mesophases. *Cryst. Growth Des.* **11**, 1193-1201.
63. Caffrey, M. (2008). On the Mechanism of Membrane Protein Crystallization in Lipidic Mesophases. *Cryst. Growth Des.* **8**, 4244-4254.
64. Liu, W., Chun, E., Thompson, A. A., Chubukov, P., Xu, F., Katritch, V., Han, G. W., Roth, C. B., Heitman, L. H., IJzerman, A. P., Cherezov, V. & Stevens, R. C. (2012). Structural Basis for Allosteric Regulation of GPCRs by Sodium Ions. *Science* **337**, 232-236.
65. Cherezov, V., Hanson, M. A., Griffith, M. T., Hilgart, M. C., Sanishvili, R., Nagarajan, V., Stepanov, S., Fischetti, R. F., Kuhn, P. & Stevens, R. C. (2009). Rastering strategy for screening and centring of microcrystal samples of human membrane proteins with a sub-10 mu m size X-ray synchrotron beam. *J. R. Soc. Interface* **6**, S587-S597.
66. Zou, Y. Z., Weis, W. I. & Kobilka, B. K. (2012). N-Terminal T4 Lysozyme Fusion Facilitates Crystallization of a G Protein Coupled Receptor. *PLoS One* **7**.
67. Zhang, C., Srinivasan, Y., Arlow, D. H., Fung, J. J., Palmer, D., Zheng, Y. W., Green, H. F., Pandey, A., Dror, R. O., Shaw, D. E., Weis, W. I., Coughlin, S. R. & Kobilka, B. K. (2012). High-resolution crystal structure of human protease-activated receptor 1. *Nature* **492**, 387-+.
68. Matthews, B. W. (1968). Solvent Content of Protein Crystals. *J. Mol. Biol.* **33**, 491-&.
69. Geertsma, E. R. & Dutzler, R. (2011). A Versatile and Efficient High-Throughput Cloning Tool for Structural Biology. *Biochemistry* **50**, 3272-3278.
70. Geertsma, E. R. (2014). FX Cloning: A Simple and Robust High-Throughput Cloning Method for Protein Expression. *DNA Cloning and Assembly Methods* **1116**, 153-164.
71. Jiang, Y. X., Lee, A., Chen, J. Y., Cadene, M., Chait, B. T. & MacKinnon, R. (2002). Crystal structure and mechanism of a calcium-gated potassium channel. *Nature* **417**, 515-522.
72. Hollenstein, K., Frei, D. C. & Locher, K. P. (2007). Structure of an ABC transporter in complex with its binding protein. *Nature* **446**, 213-216.
73. Miller, C. (2000). Ion channel surprises: Prokaryotes do it again! *Neuron* **25**, 7-9.
74. Hilf, R. J. C. & Dutzler, R. (2008). X-ray structure of a prokaryotic pentameric ligand-gated ion channel. *Nature* **452**, 375-U12.
75. Hilf, R. J. C. & Dutzler, R. (2009). Structure of a potentially open state of a proton-activated pentameric ligand-gated ion channel. *Nature* **457**, 115-U122.
76. Gerber, S., Comellas-Bigler, M., Goetz, B. A. & Locher, K. P. (2008). Structural basis of trans-inhibition in a molybdate/tungstate ABC transporter. *Science* **321**, 246-250.
77. Hohl, M., Briand, C., Grutter, M. G. & Seeger, M. A. (2012). Crystal structure of a heterodimeric ABC transporter in its inward-facing conformation. *Nat. Struct. Mol. Biol.* **19**, 395-402.
78. Kyrpides, N. C. (1999). Genomes OnLine Database (GOLD 1.0): a monitor of complete and ongoing genome projects world-wide. *Bioinformatics* **15**, 773-774.

79. Granseth, E., Seppala, S., Rapp, M., Daley, D. O. & Von Heijne, G. (2007). Membrane protein structural biology - How far can the bugs take us? *Mol. Membr. Biol.* **24**, 329-U3.
80. Tate, C. G. (2012). A crystal clear solution for determining G-protein-coupled receptor structures. *Trends Biochem. Sci.* **37**, 343-352.
81. Sarkar, C. A., Dodevski, I., Kenig, M., Dudli, S., Mohr, A., Hermans, E. & Pluckthun, A. (2008). Directed evolution of a G protein-coupled receptor for expression, stability, and binding selectivity. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 14808-14813.
82. Dodevski, I. & Pluckthun, A. (2011). Evolution of Three Human GPCRs for Higher Expression and Stability. *J. Mol. Biol.* **408**, 599-615.
83. Schlinkmann, K. M. & Pluckthun, A. (2013). Directed Evolution of G-Protein-Coupled Receptors for High Functional Expression and Detergent Stability. *G Protein Coupled Receptors: Structure* **520**, 67-97.
84. Schlinkmann, K. M., Honegger, A., Tureci, E., Robison, K. E., Lipovsek, D. & Pluckthun, A. (2012). Critical features for biosynthesis, stability, and functionality of a G protein-coupled receptor uncovered by all-versus-all mutations. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 9810-9815.
85. Scott, D. J. & Pluckthun, A. (2013). Direct Molecular Evolution of Detergent-Stable G Protein-Coupled Receptors Using Polymer Encapsulated Cells. *J. Mol. Biol.* **425**, 662-677.
86. Schlinkmann, K. M., Hillenbrand, M., Rittner, A., Kunz, M., Strohner, R. & Pluckthun, A. (2012). Maximizing Detergent Stability and Functional Expression of a GPCR by Exhaustive Recombination and Evolution. *J. Mol. Biol.* **422**, 414-428.
87. Grisshammer, R., Duckworth, R. & Henderson, R. (1993). Expression of a Rat Neurotensin Receptor in Escherichia-Coli. *Biochem. J.* **295**, 571-576.
88. Hanninen, A. L., Bamford, D. H. & Grisshammer, R. (1994). Expression in Escherichia-Coli of Rat Neurotensin Receptor Fused to Membrane-Proteins from the Membrane-Containing Bacteriophage-Prd1. *Biol. Chem. Hoppe Seyler* **375**, 833-836.
89. Tucker, J. & Grisshammer, R. (1996). Purification of a rat neurotensin receptor expressed in Escherichia coli. *Biochem. J.* **317**, 891-899.
90. Grisshammer, R., Averbeck, P. & Sohal, A. K. (1999). Improved purification of a rat neurotensin receptor expressed in Escherichia coli. *Biochem. Soc. Trans.* **27**, 899-903.
91. White, J. F., Trinh, L. B., Shiloach, J. & Grisshammer, R. (2004). Automated large-scale purification of a G protein-coupled receptor for neurotensin. *FEBS Lett.* **564**, 289-293.
92. Ho, J. T. C., White, J. F., Grisshammer, R. & Hess, S. (2008). Analysis of a G protein-coupled receptor for neurotensin by liquid chromatography-electro spray ionization-mass spectrometry. *Anal. Biochem.* **376**, 13-24.
93. Henrikson, K. P., Allen, S. H. G. & Maloy, W. L. (1979). Avidin Monomer Affinity Column for the Purification of Biotin-Containing Enzymes. *Anal. Biochem.* **94**, 366-370.

Chapter 2

A Cleavable Ligand Column for the Rapid Isolation of Large Quantities of Homogeneous and Functional Neurotensin Receptor 1 Variants from *E. coli*

Content

Chapter 2 29
2.1. Manuscript 31
2.2. Abstract 32
2.3. Introduction 33
2.4. Materials and Methods 34
2.4.1. Materials 34
2.4.2. Construct design 34
2.4.3. Expression and purification of pD-NT contructs 35
2.4.4. Coupling of pD-NT constructs to NHS-activated Sepharose 36
2.4.5. Large-scale expression and purification of agonist-bound NTR1 variants 36
2.4.6. Identification of mutant pD-NT columns for antagonist-bound receptor purifications 38
2.4.7. Large-Scale antagonist-bound TM86V purification 38
2.5. Results and Discussion 39
2.5.1. Purification strategy and ligand column design 39
2.5.2. Ligand production and coupling efficiency 40
2.5.3. Large-Scale preparation of functional NTR1 variants 41
2.5.4. Purification of antagonist-bound TM86V 44
2.5.5. Generalizability of purification principle 46
2.6. Conclusions 47
2.7. References 48

2.8. Supplementary information 50
2.9. Supplementary reference 52

2.1. Manuscript

Title: A cleavable ligand column for the rapid isolation of homogeneous and functional neurotensin receptor 1 variants from *E. coli*

Authors: Pascal Egloff[a], Mattia Deluigi[a], Philipp Heine[a], Stefanie Balada[a] and Andreas Plückthun[a]*

Affiliation: [a]Department of Biochemistry, University of Zurich, 8057 Zurich, Switzerland

* Corresponding author.
Address: Department of Biochemistry,
University of Zurich,
Winterthurerstrasse 190,
8057 Zurich, Switzerland.
Phone: +41 44 635 55 70. Fax: +41 44 635 57 12
E-mail: plueckthun@bioc.uzh.ch

2.2. Abstract

G protein-coupled receptors (GPCRs) are key players of cell signaling, thus representing important drug targets for the treatment of human diseases. Since inherent difficulties in receptor production and handling have precluded the application of many in vitro experiments, major questions about GPCR mechanisms and dynamics remain elusive to date. We recently used directed evolution in *E. coli* on neurotensin receptor 1 (NTR1) for the generation of GPCR variants with greatly elevated functional expression levels and with excellent stability in detergent micelles. In this work we outline a highly efficient purification method for our evolved receptor variants, which is based on the application of an inexpensive, disposable high-affinity ligand column as the initial purification step. The ligand resin allows isolation of correctly folded GPCR variants directly from whole *E. coli* cell lysates at the scale of 10 mg and it permits preparations of agonist- and antagonist-bound receptor samples. The purification principle presented here was key to the first structures of signaling-active NTR1 variants. Since *E. coli* is uniquely suitable for the production of fully deuterated proteins, our method provides the basis for an array of NMR experiments that were not feasible for GPCRs to date, but which will shed light on novel aspects of receptor function and dynamics.

Keywords: Membrane protein, directed evolution, G protein-coupled receptor, protein stability, ligand affinity purification

2.3. Introduction

The neurotensin receptor 1 (NTR1) is a GPCR that is expressed in the human intestine and in the nervous system [1]. It binds to the 13-amino-acid peptide neurotensin (NT), which plays important roles in hypothermia, antinociception, the pathogenesis of Parkinson's disease, schizophrenia and lung cancer progression [2-5]. Upon NT binding, the receptor triggers GDP/GTP exchange within heterotrimeric G proteins, which leads to downstream stimulation of phospholipase C and adenylyl cyclase that produce second messengers in the cytosol [6, 7]. Due to a lack of experiments tackling NTR1 dynamics, little is known about the signal-transduction mechanism across the plasma membrane, but insights of this type would improve our understanding of the receptor in its signaling network and facilitate drug development.

Like many other GPCRs, NTR1 has been studied extensively in the contexts of native membranes and in vivo [8-10]. Investigations in vitro, on the other hand, were largely precluded due to receptor instability in detergent solution. Recently, this problem was approached by alanine-scanning mutagenesis, which generated a thermostable, but signaling-deficient NTR1 variant that was fused to T4 lysozyme for crystallization in a mild lipidic cubic phase environment [11]. In parallel, we have applied directed evolution technologies that generated NTR1 variants with significantly higher stabilities [12-15]. Several of the evolved NTR1 variants could be crystallized in harsh detergent environments by standard vapor diffusion experiments, and high-resolution structures were determined in the NT-bound state, thus confirming the structural integrity of the evolved variants [16]. One of the crystallized constructs, termed TM86V-ΔIC3A, was functionally characterized. It was signaling-active, it bound agonist and antagonist with high affinities and it exhibited residual desensitizing internalization behavior typical for a GPCR.

The evolved NTR1 variants may serve as useful model GPCRs for future biophysical studies. They exhibit up to 60 fold improved functional expression levels in *E. coli* and thus benefit from several advantages of the prokaryotic expression host, such as quick genetic modification strategies, growth to high cell densities, fast doubling times, inexpensive media, absence of glycosylation and robust handling. Furthermore, prokaryotic production now allows to fully deuterate large quantities of functionally expressed NTR1 variants, which will improve signal-to-noise ratios in many NMR experiments [17] that were previously not feasible with natively produced GPCRs.

In this work, we describe a highly efficient method for the isolation of evolved NTR1 variants directly from whole *E. coli* cell extracts, which is based on ligand binding and thus allows the enrichment of correctly folded receptors only. In contrast to previously documented NTR1 preparations from *E. coli* [18-20], the strategy presented here enabled us to purify within

significantly shorter time larger quantities of receptor samples with improved homogeneity, even in very harsh short-chain glucosidic detergents. The method was key to the crystallization of several agonist-bound NTR1 variants in signaling-competent states [16], and we show that the purification concept is also applicable for the isolation of large quantities of functional receptors bound to antagonist.

2.4. Materials and Methods

2.4.1. Materials

Isopropyl-β-D-thiogalactopyranoside (IPTG) was purchased from Biosolve. All detergents were obtained from Affymetrix. Cholesteryl hemisuccinate tris salt (CHS) and lysozyme were purchased from Sigma. Empty PD10 columns, NHS-activated Sepharose, SP Sepharose, Superdex-200 10/300 and HighLoad 16/600 Superdex-200 were obtained from GE Healthcare. Complete EDTA-free protease inhibitor tablets and DNase I were purchased from Roche. The neurotensin receptor 1 (NTR1) antagonist SR142948 was obtained from Axon Medchem. Amicon Ultra concentrators were purchased from Millipore. Ampicillin was obtained from AppliChem, Ni-NTA was purchased from Qiagen. Micro Bio-spin columns were obtained from Biorad.

2.4.2. Construct design

The NT ligand constructs for NHS-activated Sepharose coupling were expressed using a pAT223-derived vector (GenBank accession number AY327138) for IPTG-inducible expression. The open reading frame encoded an N-terminally Avi-tagged protein D (pD), which is at its C-terminus connected to an internal hexa-histidine tag, followed by the linker $GS(GGGS)_4$, a human rhinovirus (HRV) 3C protease site (LEVLFQGP), two glycines and amino acids 8-13 (RRPYIL) of human/rat NT (QLYENKPRRPYIL). This construct (Fig. 2.1) is referred to as pD-NT [16]. The mutant pD-NT constructs encoded alanine substitutions in the C-terminal NT8-13 moiety (R8A, R9A, P10A, Y11A, I12A, L13A or I12A+L13A) and the HRV 3C protease site was replaced by the non-cleavable linker GGGGSGG.

All NTR1 variants were subcloned into a pBR322-derived vector, which was originally obtained as a kind gift from R. Grisshammer (National Institute of Neurological Disorders and Stroke, National Institutes of Health, Rockville USA) [12, 21]. The open reading frame of the modified vector encoded an N-terminal maltose-binding protein (MBP, including its periplasmic signal sequence) linked via a GSNN linker, hexa-histidine and an HRV 3C protease site (LEVLFQGP) to residue G50 of the receptor (sequential NTR1 numbering). The NTR1 variants were C-terminally truncated at G390 and linked via a HRV 3C protease site, a penta-asparagine linker, and a di-glycineserine linker to thioredoxin A (TrxA), which is followed by a deca-

histidine tag. Amino acids V280-I295 of the intracellular loop 3 were deleted in all illustrated purification procedures involving NTR1 variants OGG7 and HTGH4 (Fig. 2.4 C, front and middle) and in the TM86V-agonist complex purification in n-octyl-β-D-glucopyranoside (OG) (Fig. 2.4 B). For all described TM86V purifications in n-nonyl-β-D-glucopyranoside (NG), amino acids E273-T290 of the intracellular loop 3 were deleted (Fig. 2.4 C, back and Fig. 2.6).

2.4.3. Expression and purification of pD-NT contructs

Five ml 2×YT medium containing 1 % (w/v) glucose and 50 µg/ml ampicillin was inoculated with a single colony of *E. coli* BL21 harboring the pD-NT expression plasmid. The culture was incubated for 8 h at 37°C. Subsequently, 1 ml of this culture was used to inoculate 1 L pre-culture (2×YT, 1 % (w/v) glucose, 50 µg/ml ampicillin), which was grown overnight at 37°C to saturation. A 50 L fermenter (Bioengineering), containing 2×YT medium, 0.6% (w/v) glucose, and 50 µg/ml ampicillin, was inoculated to OD^{600} 0.05 and grown to an OD^{600} of 1.5 at 37°C, followed by induction with 1 mM IPTG, and growth was continued for 3 – 4 hours. The pH was kept constant at 6.5. Cells were harvested using a continuous-flow centrifuge.

All pD-NT constructs were purified at 4°C via the internal histidine-tag using immobilized metal-ion affinity chromatography (IMAC). A 100 g aliquot of the cell pellet (corresponding to less than a fifth of a fermenter yield) was resuspended by a Yellow Line DI 25 basic homogenizer (IKA) in 300 ml lysis buffer containing 50 mM HEPES, pH 8, 500 mM NaCl, 25 mM imidazole, pH 8, 30 complete EDTA-free protease inhibitor tablets, a spatula tip of DNaseI and 5 mM $MgCl_2$. The cells were lysed by one processing round of a T1.1 cell disrupter (Constant Systems) at 35,000 psi. Pelleting of cell debris was carried out in a Sorvall Evolution RC centrifuge at 30,000 *g* using an SLA1500 rotor (tilted neck 250 ml tubes, SORVALL). The supernatant was loaded onto 12 × 4 ml pre-equilibrated Ni-NTA resin in empty PD10 columns (bench-top gravity flow). The resin was washed using a Cerex SPE Processor pressure-flow device (Varian) by 10 column volumes equilibration buffer (25 mM HEPES pH 8, 500 mM NaCl, 25 mM imidazole, pH 8) and subsequently by 10 column volumes carbonate buffer (0.2 M $NaHCO_3$ pH 8.3, 500 mM NaCl, 5 mM imidazole, pH 8). The elution buffer contained 0.2 M $NaHCO_3$, pH 8.3, 500 mM NaCl and 300 mM imidazole, pH 8. The eluted protein was concentrated to approximately 20 mg/ml using 4 Amicon-15 ultra concentrators and subsequently it was dialyzed (6,000 – 8,000 Da cutoff) against 3 × 1 L coupling buffer containing 0.2 M $NaHCO_3$, pH 8.3, and 500 mM NaCl (one dialysis step was performed overnight). The purified pD-NT constructs were diluted to 12 mg/ml with coupling buffer and they were frozen in liquid N_2 in 12.5 ml aliquots (typically 5 – 6 aliquots from one purification) and stored at -80°C.

2.4.4. Coupling of pD-NT constructs to NHS-activated Sepharose

In one coupling reaction 25 ml of purified pD-NT construct at 12 mg/ml was coupled to 50 ml slurry of NHS-activated Sepharose (2 × 25 ml; #17-0906-01, GE-Healthcare). Briefly, the NHS-resin was divided equally into 10 empty PD10 columns, which were mounted to a Cerex SPE Processor pressure-flow device (Varian) at room temperature. The isopropanol-containing storage solution was drained and each column was washed with 60 ml of ice-cold 1 mM HCl. The columns were closed at the bottom prior to the addition of 2.5 ml of purified pD-NT per column. The resin was resuspended immediately in the pD-NT solution, followed by incubation for 2 h on a roller mixer (Stuart). The columns were subsequently drained and the coupling efficiency was analyzed by Bradford assays or by SDS-PAGE analysis of the drained solution (Fig. 2.2). The unreacted NHS-groups were quenched by washing each column with 10 ml of 100 mM Tris-HCl, pH 8.5 and further by resuspending the resin in 5 ml of 100 mM Tris-HCl, pH 8.5 (incubation on a roller mixer (Stuart) for 2 h at room temperature). Subsequently, 3 washing cycles using 12 ml of 100 mM sodium acetate, pH 4.5, 500 mM NaCl and 12 ml of 100 mM Tris-HCl, pH 8.5 per column and cycle were performed. To completely inactivate putatively co-purified proteases that may prevent long-term storage of the resin, each column was washed with 24 ml of 6 M GdmCl. Each column was then washed with 24 ml H_2O and with 16 ml of 20 % ethanol. The resin from all columns was collected and stored in 20 % ethanol as 50 ml slurry containing 25 ml pD-NT Sepharose (bed volume). Performing the coupling reaction on multiple benchtop-systems greatly facilitated the resuspension steps, as compared to the usage of one large column.

2.4.5. Large-scale expression and purification of agonist-bound NTR1 variants

Fermenter runs with *E. coli* BL21 Tuner cells harboring the NTR1 expression plasmids were performed as previously described [16]. In a typical purification, 50 g of cell pellet (corresponding to a 7% aliquot of a fermenter run) were resuspended by a Yellow Line DI 25 basic homogenizer (IKA) in 100 mL solubilization buffer containing 100 mM HEPES, pH 8, 20 % (v/v) glycerol and 400 mM NaCl. All subsequent steps were carried out at 4°C. 0.5 mL of 1 M $MgCl_2$, a spatula tip of DNase I, 200 mg lysozyme, 20 mL of a solution of 6 % (w/v) 3-[(3-cholamidopropyl)-dimethylammonio]-1-propane sulfonate (CHAPS) / 1.2 % (w/v) CHS, and 34 mL of 10 % (w/v) n-decyl-β-D-maltopyranoside (DM) were added to the resuspended cells while stirring. The mixture was incubated while stirring for 15 min. Sonication was performed for 30 min in an ice-water bath using a Sonifier 250 (Branson) at a duty cycle of 30 %, output 5, with sonication tip extension from Heinemann (13 mm) and a stirring bar at 250 rpm within the extraction mixture. Subsequently, 4 mL of 0.5 M EDTA was added to the extraction mixture, followed by another 30 min incubation while stirring. The suspension was transferred to one 250 ml tilted neck centrifugation tube (SORVALL) and centrifuged for 30 min at 15,000 rpm (SLA 1500 rotor). The supernatant was transferred to a 250 ml glass bottle and mixed with 5 mL

slurry pD-NT-ligand resin that was pre-equilibrated in NT wash buffer 1 (25 mM HEPES, pH 8, 10 % (v/v) glycerol, 600 mM NaCl, and 0.5 % (w/v) DM). The NT binding reaction was incubated on a roller mixer (Stuart) overnight. Subsequently, the mixture was centrifuged at 400 *g* for 10 min, and 90 % of the supernatant was decanted.

Using the remaining supernatant, the pelleted pD-NT resin was transferred to an empty PD10 column, and it was washed on a bench top in a cold room on a Cerex SPE Processor pressure-flow device (Varian) with 75 mL of wash buffer 1. The resin was subsequently washed by 40 mL of NT wash buffer 2 containing 25 mM HEPES, pH 7, 10 % (v/v) glycerol, 150 mM NaCl, 4 mM DTT and 0.4 % (w/v) NG. In the NT wash buffer 2 and in all subsequent purification buffers, various detergents at concentrations depending on their particular critical micelle concentration were tested for individual in vitro applications. For simplicity, only the one detergent (NG) is mentioned in this description (see Fig. S2.7 for a summary of alternatively applied detergents). The resin was resuspended in a small volume of wash buffer 2 within the column, containing 0.7 mg of HRV 3C protease (produced in house), followed by incubation for 2 h. The eluted protein (10 mL) was diluted three-fold with SP binding buffer, containing 10 mM HEPES, pH 7, 10 % (v/v) glycerol, 4 mM DTT, and 0.3% (w/v) NG, and it was subjected to another PD10 column (gravity flow) containing 5 mL SP Sepharose (bed volume), which had been pre-equilibrated with SP binding buffer. The SP resin was washed with 10 mL SP binding buffer, followed by 25 mL SP wash buffer containing 10 mM HEPES, pH 7.7, 10 % (v/v) glycerol, 35 mM NaCl, 4 mM DTT, and 0.3 % (w/v) NG, followed by another 3 mL SP binding buffer. Elution was carried out by ~15 mL SP elution buffer containing 10 mM HEPES, pH 7, 10 % (v/v) glycerol, 350 mM NaCl, 4 mM DTT, 0.3 % (w/v) NG and 500 nM NT1 (GPGGRRPYIL). NT1 corresponds to the C-terminal part of the fusion protein cut off by HRV 3C protease, i.e., the remaining HRV 3C protease site (GP), two linker residues (GG) and the NTR1 binding-epitope of NT, which is NT8-13 (RRPYIL).

Whenever quantitative size exclusion chromatographic analyses were performed, the NTR1 variants were concentrated by an Amicon-15 Ultra concentrator with 50 kDa cutoff to less than 500 µL (this cutoff was suitable for all tested detergents). The concentrate was transferred to an Eppendorf tube and centrifuged in a table-top centrifuge for 10 min at 10,000 *g*. The supernatant was loaded on a Superdex-200 10/300 column that was pre-equilibrated with running buffer containing 10 mM HEPES, pH 8.0, 150 mM NaCl, 4 mM DTT, 0.28 % (w/v) NG, and 100 nM NT1.

2.4.6. Identification of mutant pD-NT columns for antagonist-bound receptor purifications

The pD-NT constructs carrying various mutations in the NT moiety were expressed at the scale of 1 L *E. coli* cultures in 2×YT (0.2 % (w/v) glucose, 100 µg/ml ampicillin). After inoculation to OD^{600} of 0.05 from a saturated pre-culture, the cells were grown at 37°C and they were induced by 1 mM IPTG for 4 h after reaching OD^{600} of 0.6. The cells were harvested by centrifugation and they were resuspended in 25 ml lysis buffer (50 mM HEPES, pH 8, 500 mM NaCl, 25 mM imidazole pH 8, two tablets of complete protease inhibitor cocktail). The mutant pD-NT proteins were purified and coupled to NHS-activated Sepharose analogously to the above-mentioned protocol for the wild-type NT ligand construct, albeit at 10-fold smaller scale (using 5 ml slurry ligand column).

Small-scale TM86V purification tests were performed in parallel as follows: For all mutant pD-NT constructs (and wild-type pD-NT as control) 1 ml slurry ligand resin (equilibrated in NT wash buffer 1) was transferred to a 50 ml Falcon tube. Twenty ml of solubilized TM86V (corresponding approximately to one tenth of the preparation from 50 g of *E. coli* cells as described above) was applied to each resin and the mixtures were incubated overnight at 4°C. The pD-NT resins were subsequently pelleted at 400 *g* in a swinging-bucket rotor and transferred to empty Micro Bio-spin columns. Each resin was washed by 10 ml NT wash buffer A (25 mM HEPES, pH 8, 600 mM NaCl, 10 % (v/v) glycerol, 0.5 % (w/v) DM), followed by 3 ml NT wash buffer B (25 mM HEPES, pH 7, 2 mM DTT, 150 mM NaCl, 10 % (v/v) glycerol, 0.5 % (w/v) DM). At this stage the pD-NT resin samples corresponding to Figure 2.5 A (see below) were collected. Subsequently, 500 ul of antagonist elution buffer (10 mM HEPES, pH 7, 2 mM DTT, 150 mM NaCl, 10 % (v/v) glycerol, 0.5 % (w/v) DM, 5 mM SR142948) was added to each resin, followed by incubation of the mixture for 30 min on a roller mixer (Stuart). The columns were subsequently drained (elution: samples C in Figure 2.5) and the resins were washed with another 2 ml of antagonist elution buffer. Further samples were collected from the washed resins after elution (samples B in Figure 2.5).

2.4.7. Large-Scale antagonist-bound TM86V purification

Solubilization of TM86V, receptor immobilization at the pD-NT-P10A resin and washing of the resin by NT wash buffer 1 was performed analogously to the purification protocol for the NT-bound NTR1 variants. The ligand resin was subsequently washed by 30 ml of pD-NT-P10A wash buffer 2 (25 mM HEPES, pH 7, 10 % (v/v) glycerol, 150 mM NaCl, 4 mM DTT and 0.5 % (w/v) DM). One bed volume of antagonist elution buffer (21 mM HEPES, pH 7, 8.5 % glycerol, 128 mM NaCl, 4.3 mM antagonist, 0.85 % DM) and 500 ul of HRV 3C protease (0.7 mg) was added to the washed resin. The column was closed, the resin was resuspended and the mixture

was incubated for 2 h on a roller mixer (Stuart). Note that the antagonist stock solution for the elution buffer contained 5 mM SR142948 in 1 % (w/v) DM (the antagonist was insoluble in H_2O above 1 mM). The column was subsequently drained and further eluted by pD-NT-P10A wash buffer 2 to give a total of 10 ml elution. Subsequently, 1 ml of 10 % (w/v) NG was added to the eluted antagonist-bound TM86V and the solution was diluted by 19 ml SP binding buffer (10 mM HEPES, pH 7, 10 % (v/v) glycerol, 4 mM DTT, 0.3% (w/v) NG). All subsequent steps were performed analogously to the purification protocol for agonist-bound NTR1 variants (exception: all buffers contained 300 nM antagonist and NT1 was omitted).

2.5. Results and Discussion

2.5.1. Purification strategy and ligand column design

Many GPCR purification protocols rely on ligand-affinity chromatography to separate functional from non-functional protein. At the same time, ligand-binding usually increases receptor stability [15, 22-26]. Therefore, it is apparent that a ligand-affinity column would in principle be most effectively applied as the initial purification step in a protocol, since this would combine the benefits of an activity-based separation with an early stabilization.

Moreover, many GPCR ligands exhibit nanomolar affinities, hence significantly stronger interactions than standard histidine tags display towards nitrilotriacetic acid-immobilized nickel ions, which have micromolar affinity [27]. This can be an additional advantage during an initial purification step, particularly in the frequent case of low expression levels, when large amounts of membrane need to be solubilized and a big volume of the extraction mixture causes dilute GPCR concentrations and high quantities of competing impurities.

However, a tight-binding ligand would normally require a harsh elution step from the affinity column. This would jeopardize receptor integrity, since the receptor might denature again, and be no longer saturated with ligand, potentially eliminating any gain from such an affinity column.

In order to effectively isolate functional NTR1 variants from solubilized *E. coli* cell lysates, we generated a very inexpensive ligand-affinity resin, which allows mild receptor elution by cleaving the binary receptor-ligand complex off the column as a whole. The receptor takes the cleaved off ligand along all of the purification and as a result, it is maximally stable during the entire procedure. The ligand component consists of the minimal receptor-binding NT epitope (amino acids 8-13) fused to a fragment of protein D from phage lambda (pD) (Fig 2.1). pD is C-terminally connected to NT8-13 via a long and flexible linker, which encodes an HRV 3C protease cleavage site in close proximity to NT. Neither the linker nor NT consist of lysine residues. The construct can thus be coupled to NHS-activated Sepharose specifically via pD,

which contains 5 lysines and the free N-terminal amino group – leaving NT fully accessible for receptor binding and for proteolytic elution.

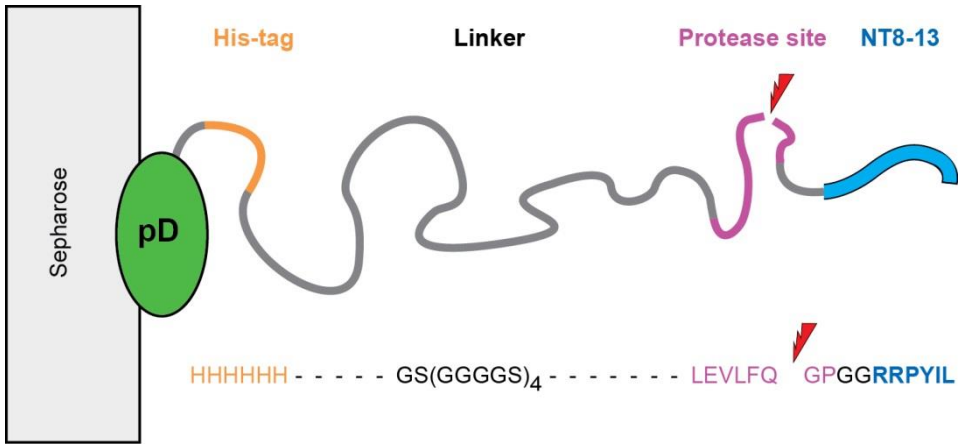


Figure 2.1 Scheme of the NTR1 ligand-affinity Sepharose resin. pD-NT contains an internal His-tag allowing its own purification in large quantities via IMAC. The ligand construct can be coupled to NHS-activated Sepharose specifically via the pD domain, as this part contains all the primary amines encoded in the construct (all the lysines and the N-terminus). The long and flexible linker between pD and the C-terminal NT8-13 (the minimal receptor-binding epitope of NT) allows maximal accessibility of the GPCR ligand. A 3C protease site in close proximity to NT8-13 permits quantitative elution of bound, functional NTR1 variants in complex with the agonist by proteolytic cleavage

2.5.2. Ligand production and coupling efficiency

Fusing the NTR1 ligand to the carrier protein pD was expected to have two beneficial effects on the ligand column production: First, N-terminal pD fusions were previously described to enhance expression levels of soluble proteins significantly in *E. coli* [28], thus permitting ligand production at large quantities. Second, peptides, such as the NT ligand, are frequently degradation-prone, when overexpressed by themselves in *E. coli* and the linkage to a folded domain was expected to minimize this problem. This is far more economical than using a synthetic peptide. Up to 100 mg pD-NT could be expressed per liter of *E. coli* culture and 700 - 900 mg were typically obtained per purification run, using standard bench-top IMAC procedures (Fig 2.2, lane 1).

The IMAC-purified pD-NT appeared to be pure on SDS-PAGE. However, it was initially not possible to achieve high coupling efficiencies to the NHS-activated Sepharose (initial yields: 40 – 70%). We suspected that small molecules containing primary amines might co-purify with pD-NT to some extent, and that these potential contaminants act as competitors in the coupling reaction. In order to prevent this, we introduced a dialysis step after the IMAC, which indeed improved the coupling efficiency to approximately 100% (Fig 2.2).

In a typical coupling reaction, 12 mg pD-NT was immobilized per milliliter of NHS-Sepharose bed volume, corresponding to a theoretical column capacity of 35 mg/ml for full-length NTR1. Ligand expression (in a fermenter), purification of a small fraction of the expressed ligand and NHS-coupling of less than half of the purified pD-NT (350 mg) required typically less than 3 days and yielded 25 ml ligand resin (bed volume), which was suitable for more than 10 large-scale GPCR purifications.

Figure 2.2 Analysis of the purification outcome and of the coupling efficiency of the ligand construct (pD-NT) to NHS-activated Sepharose. pD-NT was expressed in *E. coli*. It was purified by IMAC, followed by dialysis for the removal of small molecule contaminants containing primary amines, which appeared to act as competitors during NHS-coupling. Lane 1 shows the purity of the pD-NT construct (MW = 16 kDa) after dialysis, but prior to the coupling to NHS-activated Sepharose. Lane 2 represents a fair loading of the same solution after exposure to NHS-activated Sepharose. The absence of pD-NT in lane 2 indicates a highly efficient coupling reaction.

2.5.3. Large-Scale preparation of functional NTR1 variants

Extraction of NTR1 variants from whole *E. coli* cells was performed without prior membrane preparation using a mixture of DM, CHAPS and CHS. Subsequently, a small amount of ligand-affinity resin (2.5 ml bed volume per 200 ml solubilization reaction, cf. Materials and Methods) was applied to efficiently pull down milligram quantities of functional receptor from the soluble fraction. The ligand-affinity resin and the bound GPCRs were first washed by a DM-containing buffer and in a second step by a buffer containing a detergent of choice (bench-top column in a cold room). As observed by SDS-PAGE analysis, the receptor purity was close to 100% after this first purification step (Fig 2.4 A, lane 2 and 3). The subsequent addition of catalytic amounts of 3C protease allowed quantitative elution of the NT-bound NTR1 via ligand cleavage, and concomitantly of its cleaved fusion proteins MBP and TrxA (Fig 2.4 A, lane 5). The fusion proteins and the 3C protease were subsequently separated from the receptors by a cation-exchange step using a simple gravity-flow column format (bench-top), and the pure NTR1 variants were typically analyzed by quantitative SEC (Fig 2.4 B).

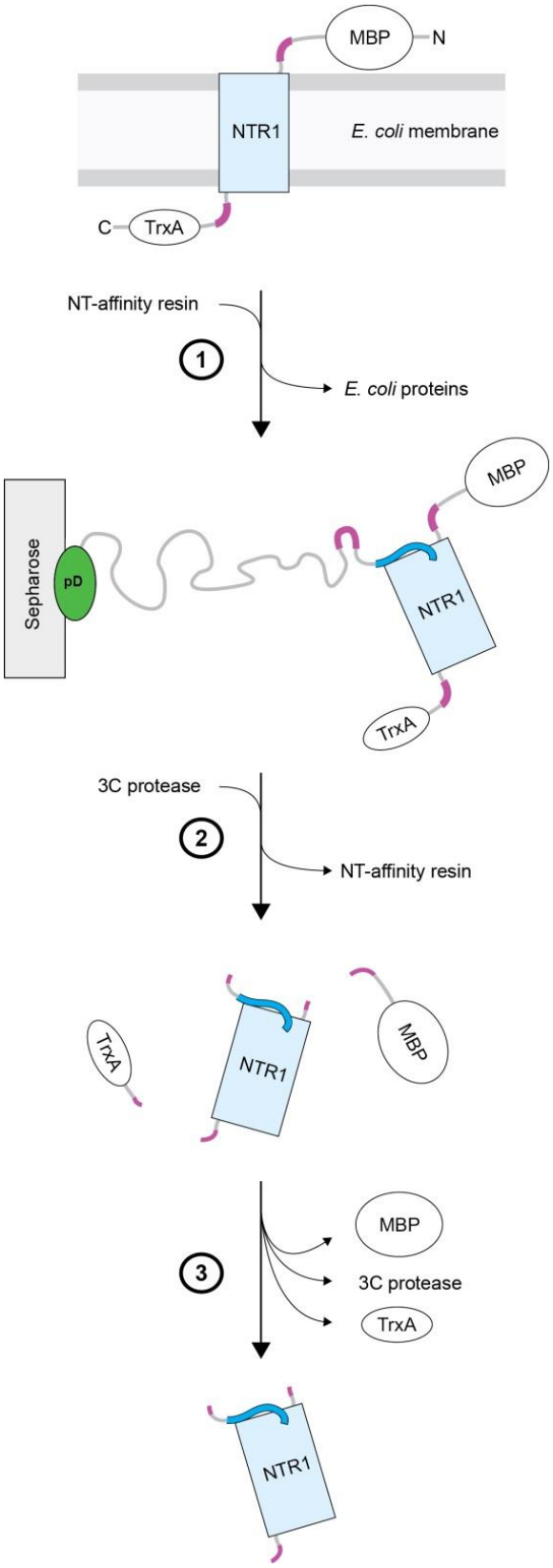


Figure 2.3 Protocol overview for the large-scale preparation of NT-bound NTR1 variants from whole *E. coli* cells. Step 1, ligand-mediated pull-down of receptors directly from solubilized *E. coli* cells and washing of the NT-affinity resin in batch. The ligand is shown in blue. Various detergents can be applied in the wash buffer at this stage, in case a detergent exchange is required. Step 2, elution of the agonist-bound receptors and of the fusion proteins from the NT-affinity resin via cleavage at the three 3C rhinovirus protease sites (magenta). Step 3, removal of the cleaved off fusion proteins by cation-exchange chromatography in batch (gravity flow). The purified receptor/NT8-13 complexes were routinely analyzed by SEC (Fig. 2.4 B and 2.4 C). This procedure is completed in less than one working day (including analytical or quantitative SEC) and allowed the isolation > 10 mg of pure receptor/agonist complexes. The ligand-mediated pull-down and the gravity-flow format allowed processing of several NTR1 variants in parallel.

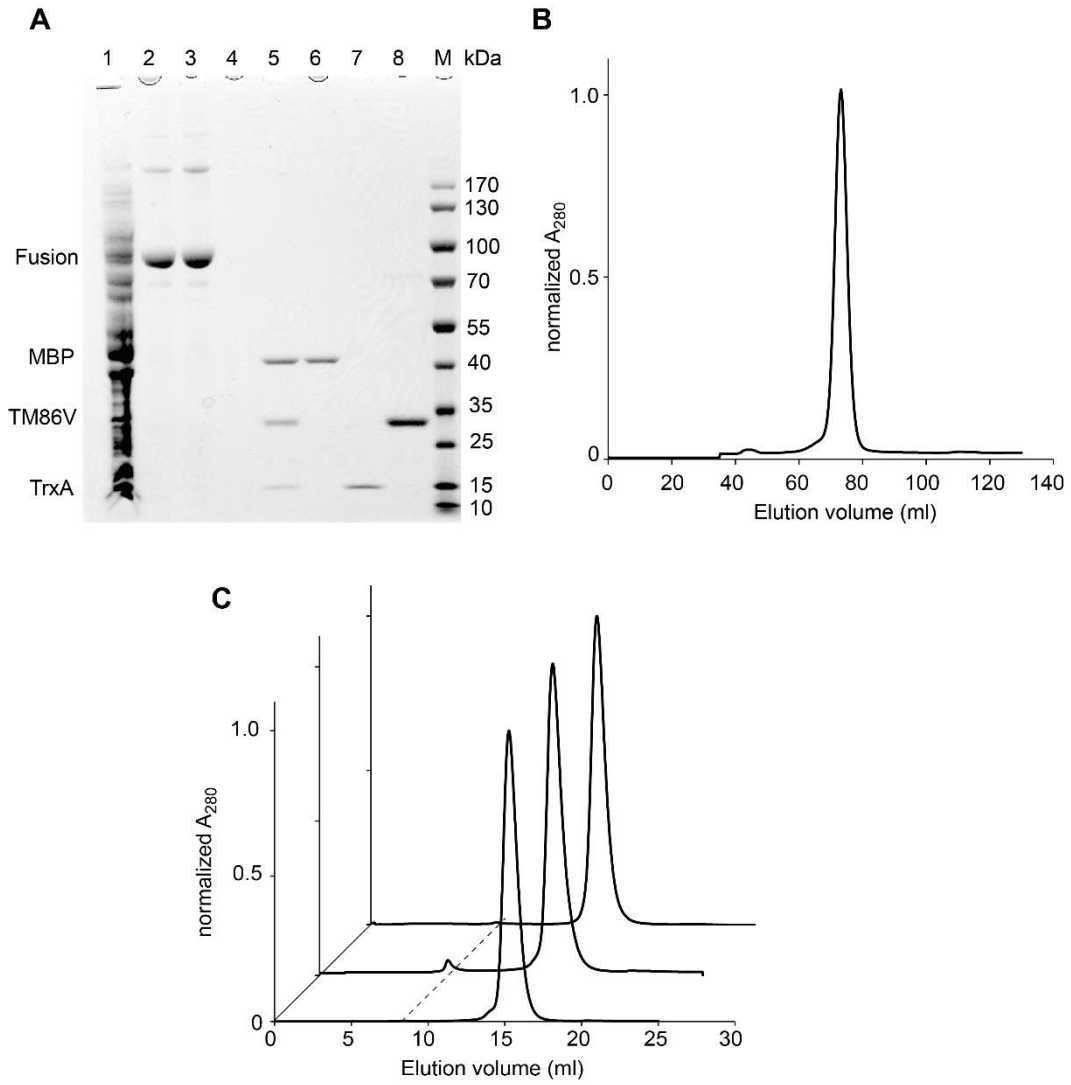


Figure 2.4 Analysis of large-scale purifications of NTR1 variants in the NT-bound state in harsh glucoside detergents. (A) SDS-PAGE analysis of a typical purification of TM86V in OG. Lane 1, DM/CHAPS/CHS-solubilized whole *E.coli* cells. Lane 2, NT-affinity resin after pull-down and wash with DM-containing buffer. Lane 3, NT-affinity resin after detergent exchange to 1% (w/v) OG. Lane 4, second wash of NT-affinity resin using OG-containing buffer (detergent exchange). Lane 5, elution from NT-affinity resin by 3C protease. Lane 6, flow-through of cation-exchange column. Lane 7, wash of cation-exchange column. Lane 8, elution from cation-exchange column. (B) Quantitative SEC (S200 HiLoad 16/600) after elution from the cation-exchange column (lane 8 of the gel) using a buffer containing 1% (w/v) OG. (C) Semi-quantitative SEC of HTGH4 (front), OGG7 (middle) and TM86V (back) in NG-containing buffer. In these cases, the detergent was exchanged to 0.3% (w/v) NG on the NT-affinity resin. The dashed line depicts the void volume of the SEC column (S200 10/300 GL).

This purification principle by ligand-mediated immobilization was successfully applied to the three evolved NTR1 variants TM86V, OGG7 and HTGH4 [15, 16, 29] (Fig 2.4 C) using several different detergents (see Fig. S2.7 for a compilation of other detergents used). Choosing the ligand-affinity step for detergent exchange was advantageous for two reasons: First, the column format allowed direct and early monitoring of NTR1 integrity in the presence of the new

detergent by UV-absorbance measurements in the wash fractions; and second, the NTR1 variants were expected to be maximally stable at this stage, due to the bound ligand, which increased the range of tolerated detergents for these intrinsically stable GPCR variants even further [15].

A typical purification of TM86V in OG yielded 10 mg of ligand-bound receptor (after SEC), which corresponds to 3.6 mg per liter of *E. coli* culture (Fig. 2.4 B). Purifications of OGG7 yielded similar amounts of agonist-bound receptors, whereas the same procedure with HTGH4 resulted typically in even higher yields (12 – 16 mg).

The described purification protocol could be completed within one working day and the simplicity of the ligand-mediated pull-down and of the gravity-flow format allowed processing of several receptor variants in parallel.

2.5.4. Purification of antagonist-bound TM86V

Given the potential of NTR1 antagonists for medical applications [30, 31], atomic-resolution structural insights about their binding mode are of high interest. We previously demonstrated that the evolved and signaling-active NTR1 variant TM86V binds to the antagonist SR142948 with high apparent affinity. TM86V was extensively characterized regarding its function and, as it yielded well diffracting crystals in complex with neurotensin, it is a promising candidate for the establishment of a purification procedure in complex with SR142948 and for subsequent crystallization trials.

In order to benefit from the ligand-mediated receptor isolation principles, we intended to adapt the agonist-based purification protocol in such a way that immobilized TM86V was no longer eluted from the pD-NT resin via ligand cleavage, but instead by competition with excess antagonist (5 mM SR142948; note, [pD-NT] < 1 µmol / ml resin). Unfortunately, even though trace amounts of antagonist-bound TM86V could be eluted using this strategy, the receptor off-rate from the pD-NT construct appeared to be too slow for quantitative TM86V preparations.

In order to identify a more suitable ligand, we tested several Sepharose resins in parallel, each displaying a different version of pD-NT with alternative mutations in the C-terminal NT. The SDS-PAGE analysis of the bound, eluted and non-eluted TM86V fractions implied that the pD-NT resins carrying the mutations P10A and I12A exhibit a reasonably reduced affinity that allows for efficient receptor binding and antagonist-mediated elution (Fig 2.5).

We then chose to implement the pD-NT-P10A resin and the antagonist competition step in the previously developed large-scale purification protocol (Fig. 2.6 A) – i.e. all other steps were performed analogously to the protocol for the receptor/agonist complex isolation described in

Fig. 2.3. This purification procedure resulted in monodisperse SEC profiles and yielded typically about 3 mg of antagonist-bound TM86V per liter expression culture, thus confirming the results of the small-scale ligand column screen and the feasibility of this purification strategy for crystallization trials (Fig 2.6 B and 2.6 C). Since the antagonist exhibits a characteristic absorbance spectrum in the UV-range, the antagonist-bound state of purified TM86V was confirmed spectroscopically (Fig. S2.8).

It is noteworthy that, in contrast to the protease-mediated pD-NT cleavage for the agonist-bound purification, elution by antagonist competition does not irreversibly remove the ligand from the column. But in order to prevent reproducibility problems and due to the simplicity of ligand resin production in large quantities, it is preferred to apply the pD-NT-P10A resin in single use as well.

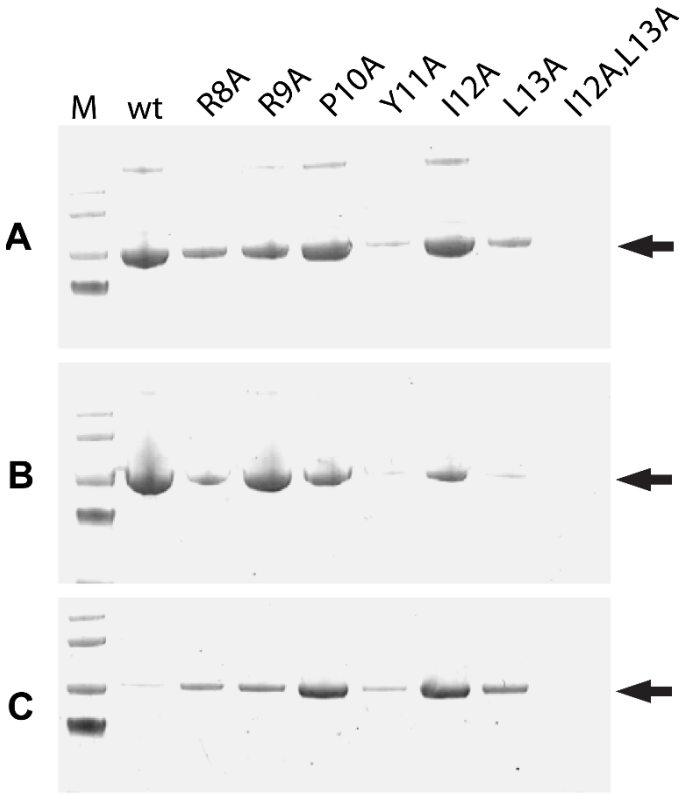


Figure 2.5 Ligand column screen guiding the identification of NT8-13 mutations that allow TM86V-binding at lower affinity and thus permit elution by competition with excess antagonist. Seven different pD-NT variants were generated in preparation of this experiment and they were coupled to NHS-activated Sepharose. Each variant encoded either a single alanine substitution within NT8-13 or the double substitution I12A / L13A. The lanes are labeled by the substituted NT8-13 amino acids of the immobilized pD-NT8-13 constructs. The figure is an SDS-PAGE analysis from the first step of small-scale TM86V purifications from solublilized whole *E. coli* cells(A) pD-NT beads after wash (removal of *E. coli* proteins). (B) pD-NT beads after antagonist-mediated elution. (C) Elution by excess antagonist. The band corresponding to TM86V is depicted by an arrow.

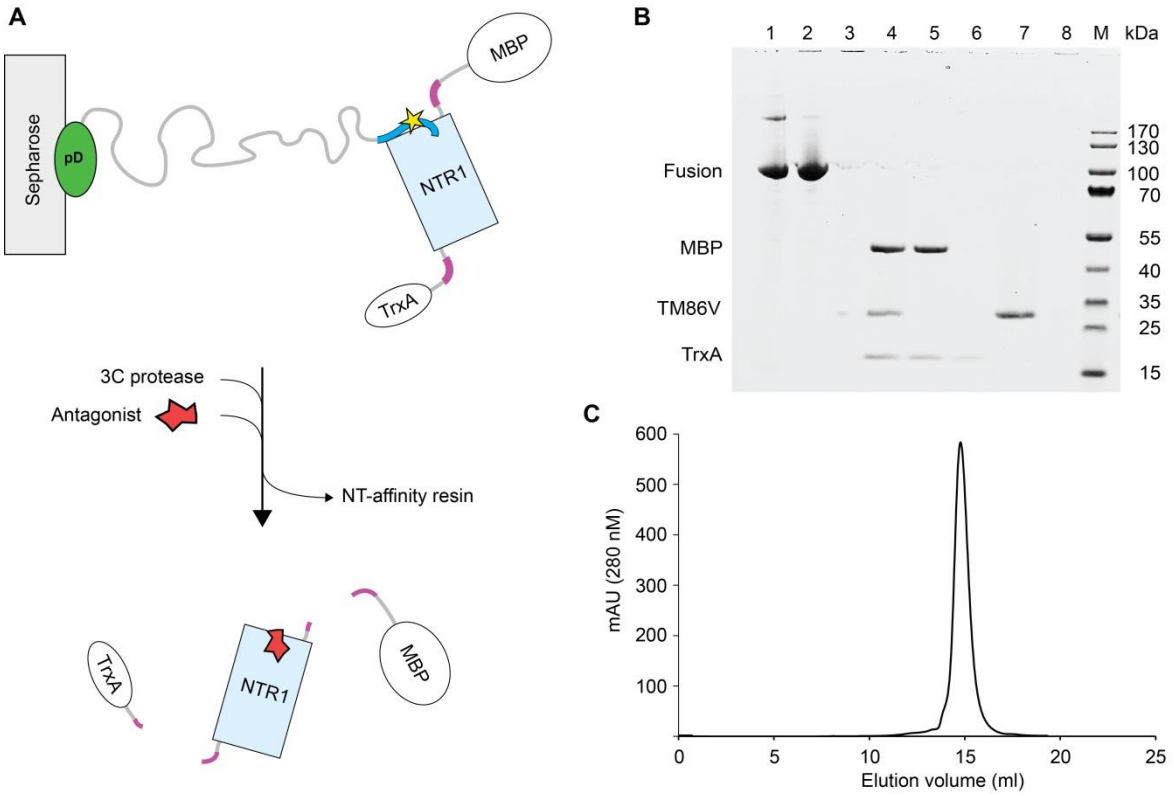


Figure 2.6 Large-scale preparation of antagonist-bound TM86V by pD-NT-P10A Sepharose. (A) Schematic representation of the key step. The ligand-affinity resin exhibits a slightly reduced interaction strength with TM86V, due to the mutation P10A (yellow star) in NT8-13 (dark blue). Quantitative elution is possible with excess of antagonist (SR142948). Note that no 3C rhinovirus protease site is encoded in the pD-NT-P10A construct. This allows simultaneous fusion protein removal and competitive elution without co-elution of free NT-P10A. The solubilization step and the removal of the cleaved fusion proteins by cation-exchange chromatography (in batch) were carried out as described for the purification of agonist-bound receptors. (B) SDS-PAGE analysis of a typical purification of antagonist-bound TM86V. Lane 1, NT-P10A-affinity resin after pull-down and wash with DM-containing buffer. Lane 2, NT-P10A-affinity resin after detergent exchange to 0.3% NG (w/v). Lane 3, NT-P10A-affinity resin after competitive elution. Lane 4, elution from NT-P10A-affinity resin by antagonist competition. Lane 5, flow-through of cation-exchange column. Lane 6, wash of cation-exchange column. Lane 7, elution from cation-exchange column. Lane 8, cation-exchange resin after elution. (C) Semi-quantitative SEC (S200 10/300 GL) of the combined elution fractions from the cation-exchange column (lane 7 in B). The complete purification was performed in 0.3% (w/v) NG after detergent exchange.

2.5.5. Generalizability of purification principle

A large number of proteins were reported to exhibit high-affinity interactions with peptides [32, 33]. Many of these ligands are devoid of lysines, hence similar economic NHS immobilization strategies by means of pD-peptide fusions may be feasible for ligand column production. In the case where a peptide-binding protein of interest benefits from ligand-mediated stabilization, or when efficient removal of aggregates is crucial, comparable purification strategies may likely improve the quantity and quality of purified proteins critically.

2.6. Conclusions

Ligand columns are very rarely used as a first step in large-scale GPCR purifications, in spite of their potential advantages, such as ligand-mediated receptor stabilization and high affinity interactions resulting in an efficient purification. There are several practical reasons that preclude the use of such columns.

Directly applying the soluble fraction to a ligand column would often prevent the repeated usage of the resin, as ligands can be unstable under these conditions (e.g. degraded by co-purified proteases) or because residual cellular debris leads to incomplete column recovery. The use of synthetic ligands can be very expensive and prohibitive for single-use columns. Also, harsh elution conditions from a tight-binding column with ensuing receptor denaturation might abrogate the effect of having captured active receptor in the first place.

However, in this work we have overcome these problems for the case of a peptide-binding GPCR by developing a high capacity ligand column, from which the receptor-ligand complex can be cleaved off. The ligand resin is so efficiently and inexpensively produced that the obtained column material is suitable for single use, thus rendering column recovery problems irrelevant.

The pD-NT Sepharose resins described in this work were the basis for time-efficient and highly reproducible receptor purifications directly from whole *E. coli* cell lysates. The outlined method was not only the foundation for the first crystal structures of signaling-active NTR1 variants expressed in a prokaryote, it will also be key to various novel in vitro studies on these receptors. As *E. coli* is the preferred expression host for isotope-labeled protein production, the way is now paved for an array of NMR studies that were not feasible using functionally expressed GPCRs to date. Future experiments based on these purification principles will likely contribute to an improved understanding of GPCR dynamics and thus facilitate drug development.

2.7. References

1. Vita, N., Laurent, P., Lefort, S., Chalon, P., Dumont, X., Kaghad, M., Gully, D., Le Fur, G., Ferrara, P. & Caput, D. (1993). Cloning and expression of a complementary DNA encoding a high affinity human neurotensin receptor. *FEBS Lett.* **317**, 139-42.
2. Bissette, G., Nemeroff, C. B., Loosen, P. T., Prange, A. J., Jr. & Lipton, M. A. (1976). Hypothermia and intolerance to cold induced by intracisternal administration of the hypothalamic peptide neurotensin. *Nature* **262**, 607-9.
3. Schimpff, R. M., Avard, C., Fenelon, G., Lhiaubet, A. M., Tenneze, L., Vidailhet, M. & Rostene, W. (2001). Increased plasma neurotensin concentrations in patients with Parkinson's disease. *J. Neurol. Neurosurg. Psychiatry* **70**, 784-6.
4. Alifano, M., Souaze, F., Dupouy, S., Camilleri-Broet, S., Younes, M., Ahmed-Zaid, S. M., Takahashi, T., Cancellieri, A., Damiani, S., Boaron, M., Broet, P., Miller, L. D., Gespach, C., Regnard, J. F. & Forgez, P. (2010). Neurotensin receptor 1 determines the outcome of non-small cell lung cancer. *Clin. Cancer. Res.* **16**, 4401-10.
5. Griebel, G. & Holsboer, F. (2012). Neuropeptide receptor ligands as drugs for psychiatric diseases: the end of the beginning? *Nat. Rev. Drug Discov.* **11**, 462-78.
6. Skrzydelski, D., Lhiaubet, A. M., Lebeau, A., Forgez, P., Yamada, M., Hermans, E., Rostene, W. & Pelaprat, D. (2003). Differential involvement of intracellular domains of the rat NTS1 neurotensin receptor in coupling to G proteins: a molecular basis for agonist-directed trafficking of receptor stimulus. *Mol. Pharmacol.* **64**, 421-9.
7. Tanaka, K., Masu, M. & Nakanishi, S. (1990). Structure and functional expression of the cloned rat neurotensin receptor. *Neuron* **4**, 847-54.
8. Dobner, P. R. (2005). Multitasking with neurotensin in the central nervous system. *Cell. Mol. Life Sci.* **62**, 1946-63.
9. Pelaprat, D. (2006). Interactions between neurotensin receptors and G proteins. *Peptides* **27**, 2476-87.
10. Hwang, J. I., Kim, D. K., Kwon, H. B., Vaudry, H. & Seong, J. Y. (2009). Phylogenetic history, pharmacological features, and signal transduction of neurotensin receptors in vertebrates. *Ann. N. Y. Acad. Sci.* **1163**, 169-78.
11. White, J. F., Noinaj, N., Shibata, Y., Love, J., Kloss, B., Xu, F., Gvozdenovic-Jeremic, J., Shah, P., Shiloach, J., Tate, C. G. & Grisshammer, R. (2012). Structure of the agonist-bound neurotensin receptor. *Nature* **490**, 508-13.
12. Sarkar, C. A., Dodevski, I., Kenig, M., Dudli, S., Mohr, A., Hermans, E. & Plückthun, A. (2008). Directed evolution of a G protein-coupled receptor for expression, stability, and binding selectivity. *Proc. Natl. Acad. Sci. USA* **105**, 14808-13.
13. Dodevski, I. & Plückthun, A. (2011). Evolution of Three Human GPCRs for Higher Expression and Stability. *J. Mol. Biol.* **408**, 599-615.
14. Schlinkmann, K. M., Honegger, A., Tureci, E., Robison, K. E., Lipovsek, D. & Plückthun, A. (2012). Critical features for biosynthesis, stability, and functionality of a G protein-coupled receptor uncovered by all-versus-all mutations. *Proc. Natl. Acad. Sci. USA* **109**, 9810-5.
15. Schlinkmann, K. M., Hillenbrand, M., Rittner, A., Kunz, M., Strohner, R. & Plückthun, A. (2012). Maximizing Detergent Stability and Functional Expression of a GPCR by Exhaustive Recombination and Evolution. *J. Mol. Biol.* **422**, 414-428.
16. Egloff, P., Hillenbrand, M., Klenk, C., Batyuk, A., Heine, P., Balada, S., Schlinkmann, K. M., Scott, D. J., Schütz, M. & Plückthun, A. (2014). Structure of signaling-competent neurotensin receptor 1 obtained by directed evolution in Escherichia coli. *Proc. Natl. Acad. Sci. USA* **111**, E655-62.
17. Venters, R. A., Farmer, B. T., Fierke, C. A. & Spicer, L. D. (1996). Characterizing the use of perdeuteration in NMR studies of large proteins C-13, N-15 and H-1 assignments of human carbonic anhydrase II. *J. Mol. Biol.* **264**, 1101-1116.
18. White, J. F., Trinh, L. B., Shiloach, J. & Grisshammer, R. (2004). Automated large-scale purification of a G protein-coupled receptor for neurotensin. *FEBS Lett.* **564**, 289-293.

19. Grisshammer, R., White, J. F., Trinh, L. B. & Shiloach, J. (2005). Large-scale expression and purification of a g-protein-coupled receptor for structure determination -- an overview. *J. Struct. Funct. Genomics* **6**, 159-63.
20. White, J. F. & Grisshammer, R. (2007). Automated large-scale purification of a recombinant g-protein-coupled neurotensin receptor. *Curr. Protoc. Protein Sci.* **Chapter 6**, 6-8.
21. Tucker, J. & Grisshammer, R. (1996). Purification of a rat neurotensin receptor expressed in Escherichia coli. *Biochem. J.* **317**, 891-899.
22. Ring, A. M., Manglik, A., Kruse, A. C., Enos, M. D., Weis, W. I., Garcia, K. C. & Kobilka, B. K. (2013). Adrenaline-activated structure of beta2-adrenoceptor stabilized by an engineered nanobody. *Nature* **502**, 575-9.
23. Ohtaki, T., Ogi, K., Masuda, Y., Mitsuoka, K., Fujiyoshi, Y., Kitada, C., Sawada, H., Onda, H. & Fujino, M. (1998). Expression, purification, and reconstitution of receptor for pituitary adenylate cyclase-activating polypeptide. large-scale purification of a functionally active G protein-coupled receptor produced in Sf9 insect cells. *J. Biol. Chem.* **273**, 15464-73.
24. Weiss, H. M. & Grisshammer, R. (2002). Purification and characterization of the human adenosine A(2a) receptor functionally expressed in Escherichia coli. *Eur. J. Biochem.* **269**, 82-92.
25. Warne, T., Chirnside, J. & Schertler, G. F. (2003). Expression and purification of truncated, non-glycosylated turkey beta-adrenergic receptors for crystallization. *Biochim. Biophys. Acta* **1610**, 133-40.
26. White, J. F. & Grisshammer, R. (2010). Stability of the neurotensin receptor NTS1 free in detergent solution and immobilized to affinity resin. *PloS One* **5**, e12579.
27. Nieba, L., Nieba-Axmann, S. E., Persson, A., Hamalainen, M., Edebratt, F., Hansson, A., Lidholm, J., Magnusson, K., Karlsson, A. F. & Plückthun, A. (1997). BIACORE analysis of histidine-tagged proteins using a chelating NTA sensor chip. *Anal. Biochem.* **252**, 217-28.
28. Forrer, P. & Jaussi, R. (1998). High-level expression of soluble heterologous proteins in the cytoplasm of Escherichia coli by fusion to the bacteriophage lambda head protein D. *Gene* **224**, 45-52.
29. Scott, D. J. & Plückthun, A. (2012). Direct molecular evolution of detergent-stable G protein-coupled receptors using polymer encapsulated cells. *J. Mol. Biol.* **425**, 662-77.
30. Rostene, W., Azzi, M., Boudin, H., Lepee, I., Souaze, F., Mendez-Ubach, M., Betancur, C. & Gully, D. (1997). Use of nonpeptide antagonists to explore the physiological roles of neurotensin - Focus on brain neurotensin/dopamine interactions. *ANN NY ACAD SCI* **814**, 125-141.
31. Moody, T. W. (2006). Peptide hormones and lung cancer. *Panminerva Med.* **48**, 19-26.
32. Okuno, Y., Yang, J., Taneishi, K., Yabuuchi, H. & Tsujimoto, G. (2006). GLIDA: GPCR-ligand database for chemical genomic drug discovery. *Nucleic Acids Res.* **34**, D673-7.
33. Reichen, C., Hansen, S. & Plückthun, A. (2013). Modular peptide binding: From a comparison of natural binders to designed armadillo repeat proteins. *J. Struct. Biol.* **185**, 147-62.

2.8. Supplementary information

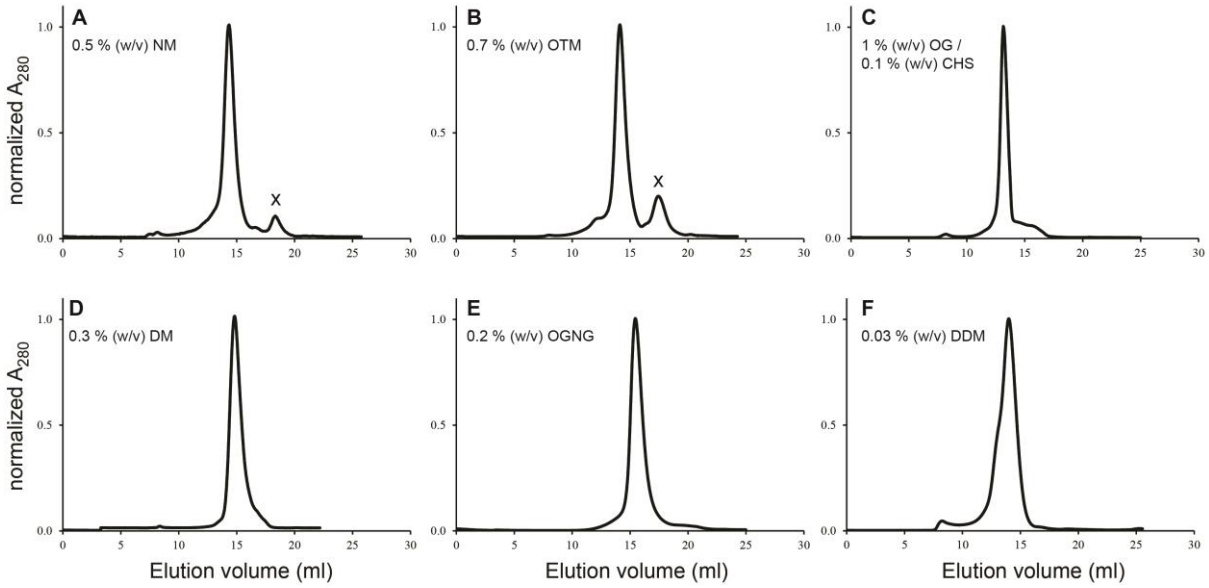


Figure S2.7 Compilation of quantitative SEC runs (S200 10/300 GL) using various detergents. This figure may guide future in vitro experiments (e.g. optimization of NMR spectra), as it provides an overview of detergent conditions that allow reasonably homogeneous preparations of evolved NTR1 variants. The chromatograms represent three different receptor variants, which were purified for various purposes (purification optimization, crystallography, NMR, MALS). All shown chromatograms represent purification procedures that were carried out using the agonist-complex purification strategy described in the main text. The exchange from DM to a detergent of choice was performed on the pD-NT ligand column and the detergent was kept constant in all subsequent buffers. Since the illustrated receptor preparations were not performed in parallel and on different FPLC systems, no conclusions from small differences in running behaviors can be drawn.
(A, B) Early purification optimization trials on NTR1-C7E02 (precursor of NTR1-TM86V [1]). "X" denotes an absorbance peak, which is due to residual amounts of TrxA, as verified by SDS-PAGE (data not shown). The cleaved fusion protein was not entirely removed during these purification attempts, due to an inefficient wash step on the SP Sepharose column. The inefficiency of TrxA removal resulted from the fact that the SP wash buffer was kept at pH 7 here, as the optimal pH of 7.7 (current protocol) was identified only at a later time point in process development. (C, D) Large-scale purification of NTR1-TM86V. (E, F) Large-scale purification of NTR1-HTGH4. Note that this variant may exhibit a tendency for dimerization under these conditions in DDM at high concentrations (shoulder at around 12.5 ml elution volume). NM: n-nonyl-β-D-maltopyranoside, OTM: n-octyl-β-D-thiomaltopyranoside, OG: n-octyl-β-D-glucopyranoside, DM: n-decyl-β-D-maltopyranoside, OGNG: octyl glucose neopentyl glycol, DDM: n-dodecyl-β-D-maltopyranoside.

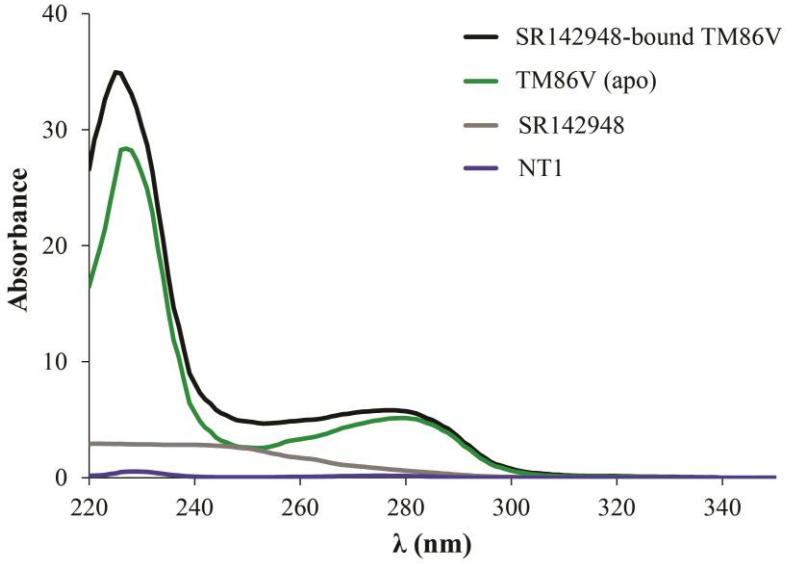


Figure S2.8 Absorbance spectra of purified TM86V, antagonist SR142948 and agonist NT1 (GGRRPYIL). All spectra were either recorded at 100 µM (SR142948) or scaled to the same concentration using estimated extinction coefficients (ProtParam online tool from EXPASY). Note that the spectrum of TM86V (apo) was determined by measuring the absorbance of purified NT1-bound TM86V followed by subtraction of the NT1 spectrum. The difference in absorbance between TM86V (apo) and SR142848-bound TM86V corresponds well to the antagonist absorbance, thus clearly confirming that the receptor can be purified in the antagonist-bound state using the described protocol. This finding is expected, since TM86V can be competed from the pD-NT-P10A column by excess antagonist (Fig 2.6) and also because competition efficiency is dependent on the position of the alanine substitution in NT8-13 (Fig 2.5).

2.9. Supplementary reference

Schlinkmann, K. M., Hillenbrand, M., Rittner, A., Kunz, M., Strohner, R. & Plückthun, A. (2012). Maximizing Detergent Stability and Functional Expression of a GPCR by Exhaustive Recombination and Evolution. *J. Mol. Biol.* 422, 414-428.

Chapter 3

Structure of Signaling-Competent Neurotensin Receptor 1 Obtained by Directed Evolution in *Escherichia Coli*

Content

Chapter 3 53

3.1. Published article 54

3.2. Supporting information 62

3.1. Published article

PNAS PLUS

Structure of signaling-competent neurotensin receptor 1 obtained by directed evolution in *Escherichia coli*

Pascal Egloff, Matthias Hillenbrand, Christoph Klenk, Alexander Batyuk, Philipp Heine, Stefanie Balada, Karola M. Schlinkmann, Daniel J. Scott, Marco Schütz, and Andreas Plückthun[1]

Department of Biochemistry, University of Zurich, 8057 Zurich, Switzerland

Edited by K. Christopher Garcia, Stanford University, Stanford, CA, and approved December 20, 2013 (received for review September 24, 2013)

Crystallography has advanced our understanding of G protein–coupled receptors, but low expression levels and instability in solution have limited structural insights to very few selected members of this large protein family. Using neurotensin receptor 1 (NTR1) as a proof of principle, we show that two directed evolution technologies that we recently developed have the potential to overcome these problems. We purified three neurotensin-bound NTR1 variants from *Escherichia coli* and determined their X-ray structures at up to 2.75 Å resolution using vapor diffusion crystallization experiments. A crystallized construct was pharmacologically characterized and exhibited ligand-dependent signaling, internalization, and wild-type–like agonist and antagonist affinities. Our structures are fully consistent with all biochemically defined ligand-contacting residues, and they represent an inactive NTR1 state at the cytosolic side. They exhibit significant differences to a previously determined NTR1 structure (Protein Data Bank ID code 4GRV) in the ligand-binding pocket and by the presence of the amphipathic helix 8. A comparison of helix 8 stability determinants between NTR1 and other crystallized G protein–coupled receptors suggests that the occupancy of the canonical position of the amphipathic helix is reduced to various extents in many receptors, and we have elucidated the sequence determinants for a stable helix 8. Our analysis also provides a structural rationale for the long-known effects of C-terminal palmitoylation reactions on G protein–coupled receptor signaling, receptor maturation, and desensitization.

membrane proteins | protein stability | protein engineering | detergents

Neurotensin is a 13-amino-acid peptide, which plays important roles in the pathogenesis of Parkinson's disease, schizophrenia, antinociception, and hypothermia and in lung cancer progression (1–4). It is expressed throughout the central nervous system and in the gut, where it binds to at least three different neurotensin receptors (NTRs). NTR1 and NTR2 are class A G protein–coupled receptors (GPCRs) (5, 6), whereas NTR3 belongs to the sortilin family. Most of the effects of neurotensin are mediated through NTR1, where the peptide acts as an agonist, leading to GDP/GTP exchange within heterotrimeric G proteins and subsequently to the activation of phospholipase C and adenylyl cyclase, which produce second messengers in the cytosol (5, 7). Activated NTR1 is rapidly phosphorylated and internalizes by a β-arrestin– and clathrin-mediated process (8), which is crucial for desensitizing the receptor (9). Several lines of evidence suggest that internalization is also linked to G protein–independent NTR1 signaling (10, 11). To improve our mechanistic understanding of NTR1 and to gain additional insight into GPCR features such as helix 8 (H8), we were interested in obtaining a structure of this receptor in a physiologically relevant state.

To date, by far the most successful strategy for GPCR structure determination requires the replacement of the intracellular loop 3 by a fusion protein, as the intracellular domain is otherwise too small to provide crystal contacts. The fusion protein approach has provided a wealth of valuable structural data on GPCRs, but as it renders the crystallized constructs signaling-inactive, the most important functionality—the activation of G proteins—cannot be confirmed for these structures. This leads inevitably to a degree of uncertainty regarding the physiological relevance of intracellular structural aspects, and it also impedes the elucidation of signaling mechanisms, as functional assays and structure determination cannot be performed with the same GPCR constructs.

Crystallization in the absence of fusion proteins was so far mainly possible for rhodopsin (12), the A_{2A} adenosine receptor (A2AR) (13), and the β_1-adrenergic receptor (14). Together, they share a high stability, which is either given naturally (rhodopsin) or it is due to stabilizing mutations. High stability appeared to be crucial for crystallographic success, as it allowed the application of harsh short-chain detergents. These tend to form small micelles, which may explain why crystal contact formation can occur under these conditions despite the small extra- and intracellular domains of class A GPCRs.

Besides the stability requirement and/or the necessity of fusion proteins, structural studies of GPCRs have also been complicated by the need of eukaryotic expression systems [e.g., *Spodoptera frugiperda* (*Sf9*) insect cells], as prokaryotes exhibit generally low functional expression levels of wild-type GPCRs. However, prokaryotes such as *Escherichia coli* offer several advantages compared with insect cells, including quick genetic modification strategies, growth to high cell densities, fast doubling times, inexpensive media, absence of glycosylation, and robust handling. Furthermore, *E. coli* is well suited for producing fully

BIOCHEMISTRY

Significance

Only a tiny fraction (<2%) of the unique structures in the protein database correspond to membrane proteins, and only a few of these are of eukaryotic origin, representing potential drug targets. The difficulties in structure determination of these proteins are due to two specific complications, which are unique for membrane proteins: first, low expression levels and, second, the necessity for detergent micelles, which are often destabilizing as they mimic the hydrophobic membrane environment only poorly. We prove that directed evolution has the potential to overcome these problems by determining several structures of evolved eukaryotic G protein–coupled receptor variants. High functional expression levels and superior receptor stability in harsh detergents allowed us to gain deeper insights into this important receptor family.

Author contributions: P.E., M.H., C.K., D.J.S., and A.P. designed research; P.E., M.H., C.K., A.B., P.H., S.B., K.M.S., D.J.S., and M.S. performed research; P.E., M.H., C.K., A.B., P.H., K.M.S., D.J.S., M.S., and A.P. analyzed data; and P.E., M.H., C.K., and A.P. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The atomic coordinates have been deposited in the Protein Data Bank, www.pdb.org [PDB ID codes 4BUO (TM86V-ΔIC3B), 3ZEV (TM86V-ΔIC3A), 4BV0 (OGG7-ΔIC3A), and 4BWB (HTGH4-ΔIC3)].

[1]To whom correspondence should be addressed. E-mail: plueckthun@bioc.uzh.ch.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1317903111/-/DCSupplemental.

isotope-labeled proteins—a crucial requirement for many NMR studies, which are limited to date.

To exploit these advantages, we recently developed a directed evolution method for high functional GPCR expression levels in *E. coli* (15). In contrast to screening a few hundred mutants one by one, this strategy allows the simultaneous, competitive testing of $>10^8$ different protein variants for highest prokaryotic expression and functionality. Briefly, diverse libraries of NTR1 variants were either obtained synthetically (16, 17) or by error-prone PCR on the wild-type sequence (15). The libraries were ligated to a plasmid encoding an inducible promoter, which was subsequently used to transform *E. coli*. Selection pressure for high functional expression levels was applied by incubating the induced cells with fluorescently labeled neurotensin, which allowed enrichment of the best expressing cells by fluorescence-activated cell sorting (FACS). The outlined procedure was performed in cycles, leading to a gradual adaptation of the NTR1 population toward high functional expression levels, and additionally, it gave rise to an increase in thermostability for certain variants.

In a second technology, termed CHESS (cellular high-throughput encapsulation, solubilization and screening), we adapted this concept to directly evolve NTR1 variants for high thermostability in short-chain detergent micelles—a property that is not only beneficial for structural studies but also for in vitro drug screening (18). The crucial development of CHESS was to surround, simultaneously, every *E. coli* cell by a semipermeable polysaccharide capsule. This allows us to solubilize the receptor mutants with harsh short-chain detergents, each mutant inside its own encapsulated cell, all at once and in the same test tube. Both the solubilized receptors and their encoding plasmids are maintained within the same capsules. Long-term incubation under these conditions followed by labeling of the encapsulated solubilized receptors with fluorescent neurotensin and rounds of FACS enrichment ensured a strong selection pressure and a gradual adaption of the NTR1 population toward high stability in harsh short-chain detergents (18).

In this work, we present the crystal structures of three evolved NTR1 variants, which were either obtained by evolving high functional expression levels in *E. coli* or by directed evolution for stability in detergent micelles. In contrast to the majority of crystallized GPCRs, our NTR1 variants are devoid of bulky modifications at the cytoplasmic face and can thus remain signaling-active, which allows us to gain unique insights into the structure–function relationship of NTR1.

Results

Directed Evolution for High Expression Levels Enabled Structure Determination of NTR1-TM86V. Directed evolution for high functional expression in *E. coli* yielded a population of well-expressed NTR1 mutants, which provided a basis for the identification of suitable variants for structural studies. We have chosen to use the variant NTR1-TM86V for crystallization, as it was the most thermostable mutant that was capable of catalyzing GDP/GTP exchange at the heterotrimeric G protein $\alpha_{i1}\beta_1\gamma_1$ (G_i) in an agonist-dependent way (see Fig. 2 *C* and *D*) (17). NTR1-TM86V harbors 11 point mutations (A86L, H103D, H105Y, A161V, R167L, R213L, V234L, I253A, H305R, F358V, and S362A) that confer the high expression levels in *E. coli* and its stability in detergent solution (Table S1). We observed that the long and putatively flexible intracellular loop 3 and the receptor termini are not required for G_i signaling and hence shortened them to aid crystallization (TM86V-ΔIC3A). The protein could be purified to homogeneity (Fig. S1) in the short-chain detergent nonyl-β-D-glucopyranoside, and it was crystallized by standard vapor diffusion techniques.

The crystal structure of TM86V-ΔIC3A at 3.26 Å [I/σ(I) = 2.0] revealed a canonical GPCR fold (Fig. 1*A* and Table S2) with seven transmembrane helices (TMs) and the prototypical amphipathic

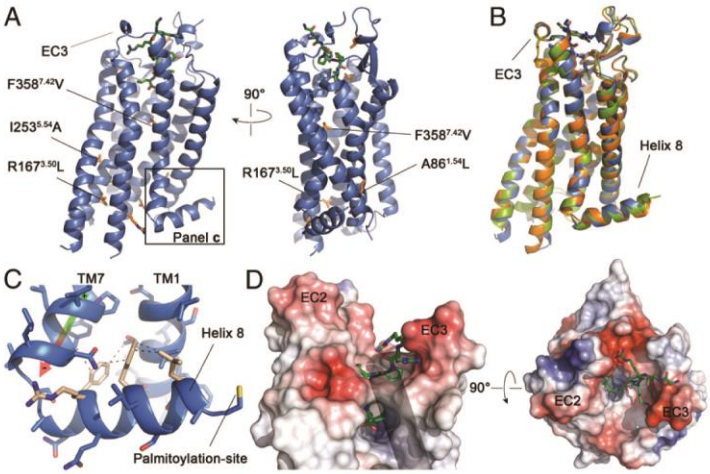


Fig. 1. Structures of three evolved NTR1 variants determined devoid of fusion proteins. (*A*) The signaling-competent NTR1-TM86V-ΔIC3A (blue) bound to its natural agonist neurotensin (green). All selected mutations for increased expression levels in *E. coli* and high stability in detergent solution are depicted (orange). (*B*) Superposition of NTR1-TM86V-ΔIC3A (blue), NTR1-OGG7-ΔIC3A (green), and NTR1-HTGH4-ΔIC3A (orange). (*C*) Close-up view of the H8 region in NTR1-TM86V-ΔIC3B. Certain hydrophobic contacts of amino acids of the semiconserved H8 motif (beige) are depicted by dashed lines for clarity. The helix dipole of TM7 is illustrated by an arrow. The first of the two palmitoylation sites adjacent to the H8 C terminus is indicated. Note the absence of the palmitoyl moiety due to the prokaryotic expression. (*D*) Vacuum-electrostatic surface representation (PYMOL) of the neurotensin-binding pocket of TM86V-ΔIC3A. Parallel (*Left*) and perpendicular (*Right*) view to the membrane. TM5 is represented as a transparent tube in the *Left* panel for clarity. Neurotensin is a 13-amino-acid peptide in vivo, but only the C-terminal residues 8–13 were reported to be relevant for binding to NTR1. Strong electron density for these six amino acids was found and allowed us to model the ligand unambiguously (Fig. S2). In addition, relatively weak electron density for two N-terminal linker amino acids (Gly–Gly) of the peptide was observed in one complex of the asymmetric unit (modeled here).

H8 (Fig. 1*C*). We observed strong electron density for the agonist neurotensin, confirming that the GPCR produced in *E. coli* reaches a functional conformation despite the absence of the eukaryotic translation and membrane insertion machinery (Fig. 1*D*, Table S3, and Fig. S2). The resolution was subsequently further improved by a change of the intracellular loop 3 deletion (TM86V-ΔIC3B), which resulted in an additional crystal contact in the same space group. TM86V-ΔIC3B was overall identical to TM86V-ΔIC3A ($RMSD_{C\alpha}$ = 0.3 Å) and could be refined to a resolution of 2.75 Å (Table S2).

TM86V-ΔIC3A Exhibits the Functional Characteristics of a Typical GPCR. To verify the physiological relevance of the initial structure of TM86V-ΔIC3A, we characterized the crystallized construct regarding ligand affinities, G protein activation, and neurotensin-dependent internalization. Ligand-binding assays on whole *E. coli* cells revealed that TM86V-ΔIC3A exhibits an apparent dissociation constant of 2.3 ± 0.4 nM for the agonist neurotensin (cf. wild-type NTR1, 2.8 ± 0.3 nM). In contrast to the agonist, the antagonist SR142948 had never been used as a ligand during directed evolution, but we still observed only a moderate increase in IC_{50} for TM86V-ΔIC3A (30 ± 2.4 nM) compared with the wild-type receptor (8.4 ± 0.9 nM) (Fig. 2 *A* and *B* and Fig. S3), which may be attributable to the point mutation F358V in NTR1-TM86V—a residue that was shown to be specifically involved in antagonist (but not agonist) binding (19).

To confirm interactions with G proteins, we measured GDP/GTP exchange in membranes containing TM86V-ΔIC3A and the reconstituted heterotrimeric G protein $\alpha_{i1}\beta_1\gamma_1$ (G_i) (Fig. 2*C* and Fig. S4). The crystallized GPCR construct exhibited a slightly increased basal GDP/GTP exchange catalysis at G_i compared with wild-type NTR1, which was further stimulated by the addition of agonist. Even though the maximal signaling level is reduced compared with wild-type NTR1, it appears that the crystallized construct is indeed able to bind to and activate G_i. To confirm these observations, we also demonstrated specific G_i binding of detergent-solubilized TM86V-ΔIC3A in a pull-down experiment using immobilized G protein on magnetic beads (Fig. 2*D*). Basal and agonist-dependent signaling of the crystallized construct TM86V-ΔIC3A was further increased by reverting the mutation $R167^{3.50}L$ [superscript according to Ballesteros–Weinstein (20)] in the highly conserved D/ERY motif (Fig. 2*C* and Fig. S5). Even though the reintroduction of $R167^{3.50}$ resulted in significantly reduced expression levels in *Sf*9 insect cells (Fig. S4*A*), the thermostability remained almost unperturbed (Fig. S4*D*).

We also investigated β-arrestin2–dependent desensitization behaviors by confocal microscopy on living HEK293T cells, which coexpressed TM86V-ΔIC3A and β-arrestin2–YFP. Despite the lacking C terminus in the crystallized construct, we observed a weak internalization when bound to fluorescent neurotensin (Fig. 2*E* and Fig. S6). Furthermore, after reconstituting $R167^{3.50}$ and the receptor C terminus, a pronounced cointernalization of β-arrestin and fluorescent neurotensin was observed, suggesting that this mutant can indeed interact with β-arrestin2 in a fashion similar to wild type (Fig. S6).

In summary, our pharmacological data clearly suggest that the crystallized NTR1 construct TM86V-ΔIC3A exhibits all essential core functions of a GPCR. Considering the simplicity of expression and genetic modification strategies in *E. coli* and the high stability of TM86V-ΔIC3A, the protein will likely serve as a valuable model system for future structural and functional studies.

Two Structures of Stability-Evolved NTR1 Variants. NTR1-TM86V was obtained by evolving high functional expression in *E. coli* and subsequently by choosing and recombining the most thermostable mutations (16, 17). In contrast, the CHESS technology can directly generate detergent-stable NTR1 variants by an evolutionary process (18). As a proof of this principle, we were interested in confirming the structural integrity of these variants as well. NTR1-OGG7 and NTR1-HTGH4 were generated by CHESS and represent the most thermostable mutants obtained so far. They crystallized readily under various conditions, and the structures were refined to 3.1 Å (OGG7-ΔIC3A) and 3.57 Å (HTGH4-ΔIC3A), respectively. Despite significant sequence variations, OGG7-ΔIC3A and HTGH4-ΔIC3A are structurally nearly identical to TM86V-ΔIC3A (TM86V-ΔIC3A/OGG7-ΔIC3A $RMSD_{C\alpha}$ = 0.4 Å; TM86V-ΔIC3A/HTGH4-ΔIC3A $RMSD_{C\alpha}$ = 0.4 Å) (Fig. 1*B* and Table S1). This suggests that the ligand-guided selection pressure has favored or preserved the same conformational state in these evolved variants, independent of the particular kind of directed evolution (for functional expression or high stability in detergents).

Because TM86V-ΔIC3A is signaling-active (Fig. 2 *C* and *D*) and exhibiting wild-type–like ligand affinities (Fig. 2 *A* and *B*) and also desensitization characteristics (Fig. 2*E* and Fig. S6), it is likely that all our structures represent a naturally occurring conformation of NTR1. Taken together, our four structures of

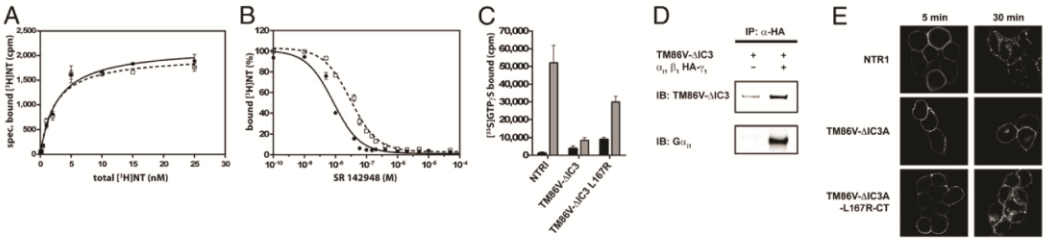


Fig. 2. Pharmacological characterizations of the crystallized NTR1 construct TM86V-ΔIC3A. (*A*) Neurotensin saturation-binding assay of wild-type NTR1 (circles) and TM86V-ΔIC3A (open squares). Note that B_{max} levels are not representative for the expression levels of the different mutants, as 10-fold more cells were used for wild-type NTR1 to obtain a similar signal-to-noise ratio—that is, the normalized B_{max} would be about 10-fold lower. (*B*) SR142948 antagonist competition binding experiment using wild-type NTR1 and TM86V-ΔIC3A. (*C*) GDP/[^{35}S]GTPγS signaling assays of wild-type NTR1, TM86V-ΔIC3A, and TM86V-ΔIC3A $L167^{3.50}R$ in insect cell membranes. Equivalent amounts of active GPCR and reconstituted G_i were assayed in the presence (gray) or absence (black) of neurotensin. The signals correspond to the average of two signaling assays performed in parallel from two independent GPCR expressions, and the error bars represent SDs. (*D*) Pull-down experiment using immobilized G_i and solubilized GPCR from *E. coli* membranes. (*E*) Confocal imaging of living HEK293T cells expressing NTR1, TM86V-ΔIC3A, or TM86V-ΔIC3A $L167^{3.50}R$-CT (reconstituted D/ERY motif and C terminus) after stimulation with fluorescent neurotensin8-13-HL647 for the indicated times.

the three different NTR1 variants exemplify that two directed evolution methods in *E. coli*, which we have recently developed, are valuable tools for structural studies of GPCRs. Our technologies have been applied successfully to a number of other receptors (21), underlining the potential of Darwinian evolution in protein research.

Improved Interhelical Surface Complementarity May Contribute to Increased Thermostability in NTR1-TM86V. We were interested in identifying the molecular causes of different thermostability characteristics among NTR1 mutants. When comparing the thermostabilities of NTR1-TM86V with one of its precursors, termed NTR1-D03 (15), we uncovered a pattern that sheds light on this issue. NTR1-D03 harbors all NTR1-TM86V mutations except A86$^{1.54}$L, I253$^{5.54}$A, and F358$^{7.42}$V. Despite only three amino acid differences, NTR1-D03 exhibited a very low thermostability in the short-chain detergent octyl-β-D-glycopyranoside, whereas NTR1-TM86V exhibited a high thermal denaturation point of 38 °C under these particularly harsh detergent conditions (Fig. 3*A*). Interestingly, the mutations cause only the replacement of hydrophobic amino acids with other hydrophobic residues. It is striking that the bulky wild-type amino acids at positions 253 and 358, where directed evolution favored a shortening of the hydrophobic side chains, would lead to obvious clashes for all common rotamers in silico (Fig. 3 *C* and *D*). At position 86, where the longer leucine was preferred over the shorter alanine, the in silico back-mutation would cause a loss of favorable van der Waals contacts between TM1 and TM2 (Fig. 3*B*). These observations suggest that improved interhelical surface complementarity contributes significantly to the high thermostability of NTR1-TM86V, and conversely, it may be speculated that optimal helix packing is not required for this particular state of wild-type NTR1 in nature.

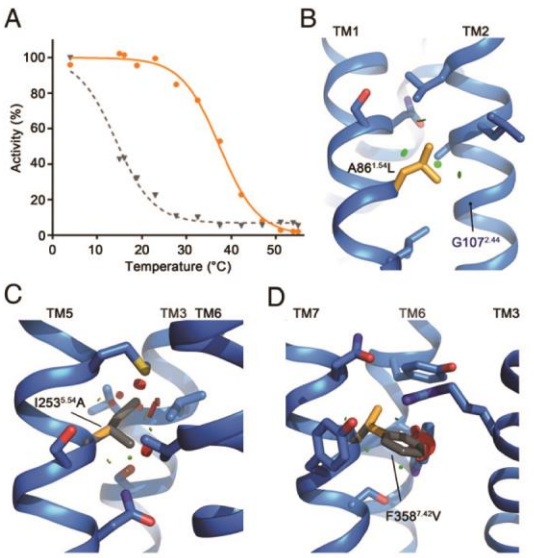


Fig. 3. Improved interhelical surface complementarity correlates with increased thermostability. (*A*) Thermostability assays of NTR1-D03 (gray) and NTR1-TM86V (orange) bound to neurotensin in the harsh detergent octyl-β-D-glucopyranoside. Note that the low stability of NTR1-D03 in this detergent did not permit an accurate determination of its thermal denaturation transition point. NTR1-D03 and NTR1-TM86V are identical except for three additional mutations in NTR1-TM86V, which must confer this thermostability difference. (*B–D*) The structure of TM86V-ΔIC3B illustrates the 3-dimensional context at these positions. In silico back-mutating the selected residues (orange) to the wild-type amino acids (gray) would either cause a reduction of favorable van der Waals contacts (green circles in *B*), or it would lead to steric clashes (red circles in *C* and *D*). For the wild-type residues in *C* and *D*, the most common rotamers based on the library of PYMOL are shown. (See Fig. S8 for additional rotamers.)

NTR1 Can Adopt a Prototypical Inactive State at the Cytosolic Domain. In activated GPCR states, the cytosolic ends of TM5 and TM6 were described to be tilted outward relative to their inactive state (22). This is observed in the most prominent way in the structure of the β_2-adrenergic receptor bound to $G\alpha_s\beta_1\gamma_2$ (23). Even though TM86V-ΔIC3A is bound to its natural agonist and capable of triggering GDP/GTP exchange at G_i, the conformations of TM5 and TM6 that were trapped in the crystal are highly similar to dark-state rhodopsin, which represents an inactive or "closed" state (Fig. 4, Fig. S5, and Fig. S7). Our finding is in agreement with other agonist-bound GPCR structures that were crystallized in inactive states, and it provides further evidence that fully active states require the G protein for stabilization. Several structural studies on rhodopsin and on the β_2-adrenergic receptor suggest that the observed closed conformation would occlude the G protein–binding site (22–25). Nevertheless, TM86V-ΔIC3A is able to functionally couple to G proteins to a certain degree (Fig. 2*C*), suggesting that the crystallized construct exhibits structural flexibility and allows a conformational change when bound to agonist. The evolved NTR1 thus shows characteristics consistent with a conformational equilibrium typical for GPCRs: In the absence of a G protein, energetically the most favorable arrangement at the intracellular side of TM86V-ΔIC3A is likely the inactive conformation that was trapped in the crystal. This may also be the case for wild-type NTR1 in the apo-state, as it exhibits very low basal signaling activity toward G_i (Fig. 2*C*).

Structural Comparison of the Evolved NTR1 Variants to NTR1-GW5. The observation of a prototypical inactive NTR1 state represents one of the unique features that distinguishes the structures presented in this work from the structure of the NTR1 variant GW5 [Protein Data Bank (PDB) ID code 4GRV] (Fig. 5*A*) (26). The mutations present in NTR1-GW5 were identified by alanine-scanning mutagenesis, and the protein required expression in *Sf*9 insect cells, fusion to T4 lysozyme replacing intracellular loop 3 (GW5-T4L), and crystallization in the presence of ligand in lipidic cubic phase (26). The crystallized construct GW5-T4L exhibits a 200-fold increased IC_{50} value for the antagonist SR48692, and it is signaling-inactive in the presence and absence of the fusion protein. The authors suggested nevertheless that the structure represents an active-like conformation, based on a partial outward tilt of the intracellular end of TM6 and on the observation of a hydrogen bond between R167$^{3.50}$ and N257$^{5.58}$. Similar features had previously been found in other GPCR structures, which represent most likely active states (27).

White et al. (26) suggested that an unusual elongation of TM7 may cause the observed lack of signaling. Indeed, a comparison with our structures reveals that TM7 is extended in GW5-T4L by a peptide segment that corresponds to the amphipathic H8 (Fig. 5 *A* and *B*). A canonical H8 would clash into a neighboring molecule in the lipidic cubic phase crystal lattice of GW5-T4L. Instead, the H8 segment resides at the center of the cytosolic domain, where it blocks the prototypical inactive position of TM6 and thus also the putative G protein–binding pocket—an arrangement that has never been observed for other GPCRs.

In contrast, all our structures suggest a canonical H8 (Fig. 1*B*), and one of the two TM86V-ΔIC3A molecules in the asymmetric unit exhibits no crystal contacts at H8. Furthermore, as described above, TM6 is positioned in a prototypical inactive conformation when bound to agonist. The observed outward tilt of TM6 in

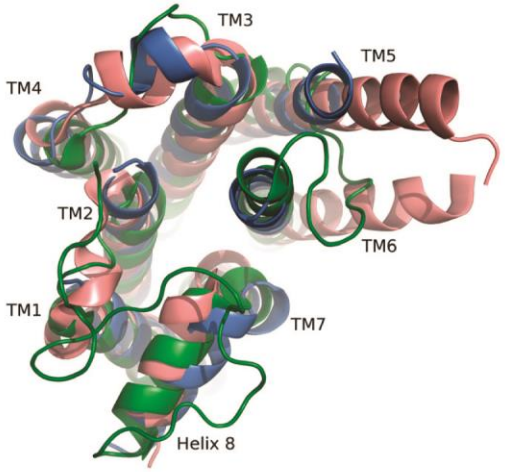


Fig. 4. View from the cytosol onto the superposition of TM86V-ΔIC3A (blue), dark-state bovine rhodopsin (green, PDB ID code 1U19), and β_2-adrenergic receptor bound to $G\alpha_s\beta_1\gamma_2$ (salmon, PDB ID code 3SN6; $G\alpha_s\beta_1\gamma_2$ is omitted).

GW5-T4L could thus alternatively be explained by the unusual contacts of TM7 to TM6.

Although the cytosolic regions of the evolved NTR1 variants described here are very different from GW5-T4L, at the extracellular side, only one major discrepancy can be observed (Fig. 5 *C* and *D*). The $2F_O$-F_C omit map of TM86V-ΔIC3B suggests a single α-helical turn of ECL3 with several ligand contacts (including a salt bridge between D336 and R9 of neurotensin). The same arrangement was found in all structures of the evolved mutants, and it is in agreement with published mutagenesis data (28, 29). In GW5-T4L, on the other hand, the loop contains no secondary structural element and it was modeled significantly more distant to neurotensin with an unusual *cis*-peptide bond following D336 (Fig. 5*D*).

NTR1-Specific Determinants of Reduced H8 Stability. Practically all high-resolution GPCR structures exhibited an amphipathic H8 following TM7. Its presence thus appeared to be a general feature of GPCRs, but surprisingly, the recently determined structures of proteinase-activated receptor 1 (PAR1), chemokine receptor 4 (CXCR4), and NTR1 (GW5-T4L) do not exhibit H8. How relevant are these findings physiologically?

Class A GPCRs exhibit relatively small intracellular domains, and it is thus apparent that the presence or absence of the complete H8 is a major factor determining the characteristics of the cytosolic interface. Multiple lines of evidence suggest an important functional role for this protein segment including G protein coupling and β-arrestin activation (30–36). Our finding of a canonical H8 in the evolved NTR1 variants now shows that a GPCR, which was previously crystallized without H8 being formed, can exhibit a canonical H8 structure (Fig. 1*C*).

To understand structural key features that are critical for the presence (or absence) of the canonical H8, we compared our structures to A2AR (PDB ID code 4EIY) (Fig. 6 *A* and *B*): A2AR likely exemplifies one of the most stable H8 arrangements, as the amphipathic helix shows large contacts to TM1, TM2, IC1, and TM7 and because all reported A2AR structures exhibited a canonical H8 irrespective of the presence or absence of crystal contacts in this region and despite a variety of crystallization conditions.

Both A2AR and NTR1 encode the semiconserved H8 motif F(R/K)xx(F/L)xxx(L/F) (Fig. 7). A common feature of our structures and of A2AR is the location of the positively charged guanidinium group of $R^{8.51}$ (Ballesteros–Weinstein numbering, 8.50 = F376 in NTR1) at the negative helix-dipole at the C terminus of TM7 (Fig. 6 *A* and *B*). This interaction likely contributes to the stabilization of the helix break between TM7 and H8, which is not encoded per se, as helix-destabilizing residues (prolines or glycines) are absent in the peptide segment connecting the two helices. Nonetheless, $R(K)^{8.51}$ is conserved among class A GPCRs (81%) and similar interactions can be found in the majority of published GPCR structures. Another similarity between the crystal structures of A2AR and NTR1 is the absence of palmitoylation membrane anchors, in the case of A2AR because of the absence of cysteines adjacent to the H8 C terminus and in the case of NTR1-TM86V because of expression in a prokaryotic system.

Clearly distinct interactions in NTR1 and A2AR are observed for the most conserved residue $F^{8.50}$ of the H8 motif. In A2AR, $F^{8.50}$ is entirely surrounded by a hydrophobic pocket (Fig. 6*A*). Most prominently, the conserved $Y^{7.53}$ of the NPxxY motif at H7 exhibits π–π stacking interactions to $F^{8.50}$, typical for a receptor in the inactive state (13). On the opposite side of its aromatic side chain, $F^{8.50}$ engages in van der Waals contacts to the poorly conserved L37 in IC1 (sequential numbering of A2AR used for this residue). The aromatic ring of $F^{8.50}$ is thus sandwiched between $Y^{7.53}$ and L37. Furthermore, several hydrophobic residues

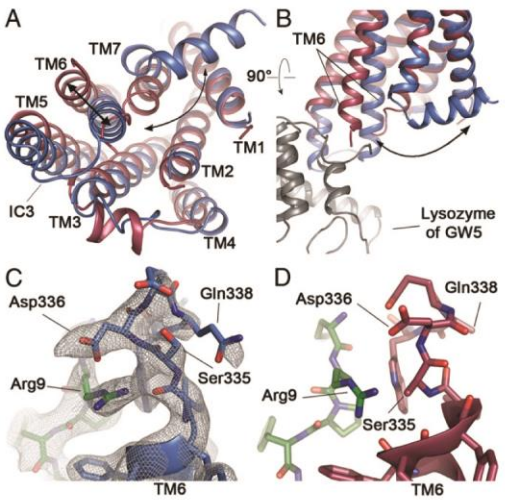


Fig. 5. Comparison of neurotensin-bound TM86V-ΔIC3B and GW5-T4L. (*A*) Superposition of TM86V-ΔIC3B (blue) and GW5-T4L (red), view from the intracellular side. The fused T4 lysozyme of GW5 replacing IC3 is omitted for clarity. Black arrows highlight the two different C-terminal conformations and the alternative states of TM6. (*B*) View along the inner leaflet of the membrane, including a part of the fused T4 lysozyme of GW5. (*C* and *D*) Comparison of the ligand-binding pockets, focusing on the interactions of EC3 with neurotensin (green). (*C*) The $2F_O$-F_C omit map of TM86V-ΔIC3B (contoured at a σ level of 1.2) suggests a single α-helical turn of ECL3 in close proximity to the ligand. (*D*) In GW5-T4L (26) the loop contains no secondary structural element and it was modeled more distant to the peptide agonist with a *cis*-peptide following Asp336. Side chains of Ser335 and Gln338 were modeled up to C_β only.

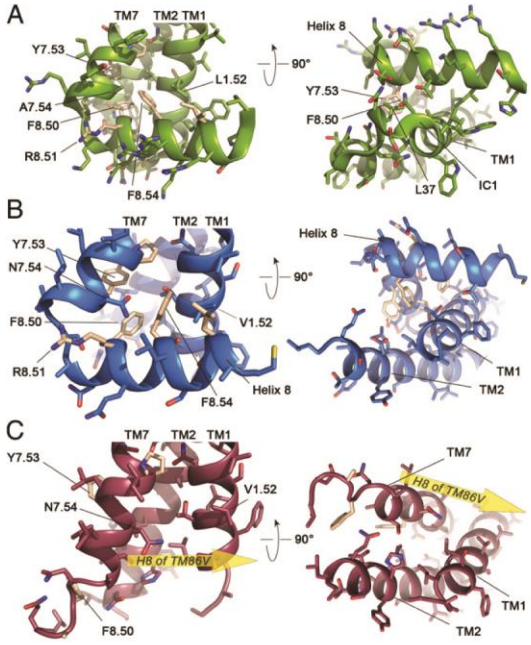


Fig. 6. Key interactions of the H8 region in A2AR and NTR1. (*A–C*) Depicted are the cytosolic ends of TM1, TM2, TM7, and H8 of A2AR (*A*; PDB ID code 4EIY), TM86V-ΔIC3B (*B*), and GW5-T4L (*C*; PDB ID code 4GRV) viewed parallel to the membrane (*Left*) or from the intracellular side (*Right*). The yellow arrow in the GW5-T4L structure corresponds to the approximate position of H8 in TM86V-ΔIC3B.

of TM1, TM2, TM7, and H8 contact the CH groups of the aromatic system of $F^{8.50}$ (not depicted in Fig. 7), and in addition, $F^{8.54}$ of the H8 motif covers the hydrophobic pocket of $F^{8.50}$: $F^{8.54}$ is accommodated between the poorly conserved $A^{7.54}$ and $L^{1.52}$.

In our NTR1 structures, $F^{8.50}$ mediates only weak interhelical interactions and no π–π stacking to $Y^{7.53}$, as it is only partially inserted into the pocket between TM1, TM2, and TM7 (Fig. 6*B*). This is due to the following three reasons: First, $F^{8.54}$ fails to cover the pocket, because it cannot obtain an analogous rotamer conformation to A2AR; it would clash into the longer side chain of $N^{7.54}$ ($A^{7.54}$ in A2AR), and $V^{1.52}$ would not be long enough to stabilize the A2AR-like rotamer ($L^{1.52}$ in A2AR). Second, the pocket itself is considerably different, as the weakly conserved residues at TM1, TM2, and TM7 contacting the CH groups of the aromatic system of $F^{8.50}$ in A2AR are less hydrophobic in NTR1 (not depicted in Fig. 6*B*). And third, because IC1 of NTR1 is longer and presumably flexible (disordered in all our structures and in GW5-T4L), it does not provide a hydrophobic residue like L37 in A2AR to sandwich $F^{8.50}$ from the intracellular side (Fig. 6 *B* and *C*). Additionally, the absence of a structured IC1 in NTR1 causes also another lack of interactions to H8, as the loop mediates not only the L37 to $F^{8.50}$ contacts in A2AR but also extensive interactions with nonconserved H8 residues (Fig. 6*A*).

In summary, the shape complementarity (37) between H8 and the receptor is significantly worse in NTR1 (S_c = 0.642) than in A2AR (S_c = 0.81) and the buried surface area is strongly reduced (NTR1, 222 Å^2; A2AR, 303 Å^2). As described above, these differences are due to alternative amino acids at poorly conserved positions in IC1, TM1, TM2, and TM7, including the residues constituting the pocket of $F^{8.50}$, and they imply that the canonical H8 arrangement in NTR1 is significantly less stable than in A2AR.

Discussion

Ligand and Palmitoylation Dependence of the Canonical H8 State. In this work, we present the agonist-bound structures of the three NTR1 variants TM86V, OGG7, and HTGH4, which were generated by directed evolution for high functional expression and for stability in short-chain detergents. In contrast to most other crystallized GPCR constructs so far, TM86V-ΔIC3A not only exhibited wild-type–like ligand-binding properties, it was also able to signal to G_i to some extent. Moreover, when expressed in eukaryotic cells, the classical features of receptor desensitization and internalization were detected, suggesting that the structure derived from this construct resembles a physiologically relevant state. The canonical inactive-like positioning of TM6 is distinct from the outward tilted helix in GW5-T4L, and it is only permitted because of the presence of a canonical H8 that does not occlude this position and the putative G protein–binding pocket. We observed elevated B-factors in the H8 region for all our structures (Fig. S9), and in addition, we found comparatively weak contacts to TM1 and TM7 and a lack of interactions to IC1 and TM2 (Fig. 6*B*). Considering these observations and the absence of H8 in GW5-T4L, it is tempting to speculate that the canonical H8 of NTR1 is of lower stability than that of the prototypical version in A2AR and/or only partially occupied under certain conditions.

PAR1 and CXCR4 may represent even more extreme cases in this regard, as none of their crystal structures exhibited H8 (38, 39). These absences can be explained by the fact that they are not only different at the nonconserved positions, which cause the reduced H8 stability in NTR, but also by the observation that these two receptors lack parts of the rather conserved H8 motif (Fig. 7). This correlation points to a sequence-specific origin, and thus a naturally evolved feature of physiological relevance.

	TM7								H8																
	7.49	7.50	7.51	7.52	7.53	7.54	7.55	7.56	8.50	8.51	8.52	8.53	8.54	8.55	8.56	8.57	8.58	8.59	8.60	8.61	8.62	8.63	8.64
A_2AR	N	P	F	I	Y	A	Y	R	I	R	E	F	R	Q	T	F	R	K	I	I	R	S	H	V	L	R
NTR1	N	P	I	L	Y	N	L	V	S	A	N	F	R	Q	V	F	L	S	T	L	A	C	L	C	P	G
PAR1	D	P	L	I	Y	Y	Y	A	S	S	E	C	Q	R	Y	V	Y	S	I	L	C	C	K	E	S	S
CXCR4	N	P	I	L	Y	A	F	L	G	A	K	F	K	T	S	A	Q	H	A	L	T	S	V	S	R	G
	N	*P*	*X*	*X*	*Y*							*F*	*R*	*X*	*X*	*F*	*X*	*X*	*X*	*L*						

Fig. 7. Sequence alignment representing the end of TM7 and H8. The sequences are numbered according to Ballesteros–Weinstein (residue 8.50 chosen as F376 of NTR1). The NPxxY and F(R/K)xx(F/L)xxx(L/F) motifs are highlighted (green) and putative palmitoylation sites [experimentally confirmed in NTR1 (44, 49)] are depicted (yellow).

H8 dynamics have previously been investigated for a number of GPCRs (34, 40, 41). Among all published GPCR structures that include the amphipathic helix, it can be observed that alterations of poorly conserved residues cause a variety of subtly deviating H8 arrangements. The resulting stability differences of H8 likely reflect an evolutionary adaptation of every receptor to the particular requirement on its amphipathic helix. Importantly, the most conserved interaction (the stacking of $F^{8.50}$ and $Y^{7.53}$) was described to be disrupted upon agonist binding, as $Y^{7.53}$ flips toward the G protein–binding cavity upon receptor activation (42). The current body of high-resolution structural data therefore suggests that a reduction of forces that keeps H8 in its canonical arrangement is a common theme of GPCR activation.

Besides intramolecular interactions, another important parameter influences the stability of the amphipathic helix: the number and positions of palmitoyl anchors at its C terminus. Palmitoylation is known to be a reversible and dynamic protein modification that can be cell-cycle–dependent (43) and developmentally regulated (44). In GPCRs, the palmitoylation state was reported to affect G protein signaling, receptor maturation, membrane delivery, phosphorylation efficiency, and desensitization (44, 45). Our finding of an unstable canonical H8 in nonpalmitoylated NTR1 implies that the occupancy of the canonical H8 state of this receptor—and potentially also of other GPCRs—may depend crucially on the palmitoylation state. Considering that the presence or absence of H8 certainly represents an important source of binding specificity to cytosolic interaction partners, it is conceivable that palmitoylation/depalmitoylation events exert their physiological effects in many cases via modulating the stability and dynamics of H8.

Potential of Directed Evolution for Membrane Protein Research. Most membrane proteins are unsuitable for high-resolution structure determination, because of difficulties in overexpression, instability in detergent solution, or both. To date, the most successful approaches to circumvent these problems rely on trial-and-error procedures, like homology screens or alanine scans, which involve expression, stability, and purification tests of individual proteins in high-throughput formats. Miniaturization has indeed advanced membrane protein structural biology significantly in recent years, but given the resources it takes and the still striking underrepresentation of structural data in the PDB, it is apparent that alternative approaches are needed.

Loss of functionality and low sequence identity to the protein of interest (e.g., by using a bacterial homolog) are frequently accepted as necessary evils on the way to the structure. We have shown in this work that a fundamentally different approach was successfully applied to generate several crystallizable GPCR variants with high sequence identity to the protein of interest (93.2–97.5%) (15–18). Instead of screening mutants or homologs one by one, our method exploited the power of evolution on populations of more than a hundred million GPCR variants at once. Analogous to natural evolution, directed evolution amplified favorable GPCR traits through the alternation of random mutagenesis and selection pressure, allowing a gradual adaptation of the characteristics of the whole GPCR population toward the selected phenotype—it tailored an array of GPCR variants with suitable properties for structural biology independent of previous structural knowledge.

Importantly, the evolutionary system allowed us to determine structures of GPCRs produced in *E. coli*, thus establishing a prokaryote as a novel and robust host for quantitative, functional, and very rapid GPCR overexpression (15, 21). As *E. coli* is well suited for producing isotope-labeled proteins, we also provide the basis for an array of NMR studies that were not feasible for this class of membrane proteins so far. Furthermore, the high stability of functional GPCRs generated by directed evolution will facilitate high-throughput ligand screening in vitro, and thus likely contribute to the discovery of new drugs.

Materials and Methods

Construct Design and Expression for Crystallization. All NTR1 variants were expressed in *E. coli* using an isopropyl-β-D-thiogalactopyranoside–inducible pBR322-derived vector, which was derived from a plasmid originally obtained as a kind gift from R. Grisshammer (National Institute of Neurological Disorders and Stroke, National Institutes of Health, Rockville, MD) (46–48). They were *N*- and C-terminally truncated at G50 and G390, respectively, and linked via human rhinovirus 3C protease sites to maltose-binding protein (N terminal) and thioredoxin (C terminal). Amino acids V280-I295 were deleted in the constructs ΔIC3A and E273-T290 in ΔIC3B. Directed evolution of NTR1 was performed as previously described (17, 18). Full details are given in *SI Text*, and a list of all evolved mutations is given in Table S1.

Purification and Crystallization. Whole *E. coli* cells were solubilized in 50 mM Hepes pH 8, 10% (vol/vol) glycerol, 200 mM NaCl, protease inhibitor tablets (Roche), and 0.6% 3-[(3-cholamidopropyl)-dimethylammonio]-1-propane sulfonate (CHAPS), 0.12% cholesteryl hemisuccinate tris salt (CHS), and 1.6% (wt/vol) decyl-β-D-maltopyranoside. All NTR1 variants were purified based on ligand affinity, cation exchange, and size exclusion in nonyl-β-D-glucopyranoside, and they were crystallized in standard vapor diffusion experiments using various mixtures of glucoside detergents and cholesterol hemisuccinate as additives. (See *SI Text* for details.) The reservoir solutions of the different NTR1 crystallization conditions varied significantly regarding buffering compound (acetate pH 5.5 or glycine pH 9.4), salt (500 mM or 2 M NaCl or 0.2 M $CaCl_2$), and PEG 600 concentrations [20% (vol/vol)–26% (vol/vol)]. (See *SI Text* for details.) Diffraction data were collected from one single crystal per protein at the Swiss Light Source, and the structure was determined by molecular replacement. (See *SI Text* for details.)

Functional Assays. Ligand affinity measurements were performed on whole *E. coli* cells using either ^{3}H-neurotensin or ^{3}H-neurotensin and the unlabeled NTR1 antagonist SR142948 for competition experiments. (See *SI Text* for details.) Signaling assays were performed with purified G_i protein (expression in *Sf9* insect cells) composed of $G\alpha_{i1}$, $G\beta_1$, and $G\gamma_1$ and a defined amount of active NTR1, TM86V-ΔIC3A, or TM86V-ΔIC3A-$L167^{3.50}$R on urea-washed membranes. (See *SI Text* for details.) Pull-down experiments were performed with purified G_i and solubilized *E. coli* membranes containing the expressed TM86V-ΔIC3A. (See *SI Text* for details.) Fluorescence microscopy was performed on living HEK293T cells that were transiently transfected with NTR1 variants and β-arrestin2–YFP. (See *SI Text* for details.)

ACKNOWLEDGMENTS. We thank Prof. Raimund Dutzler for comments on crystallization procedures, Peer Mittl for support during structure determination, Mattia De Luigi and Christian Schori for insights regarding GPCR and G protein purification, and Beat Blattmann and Céline Stutz [National Center of Competence in Research (NCCR) crystallization facility] for their efforts during initial crystallization screening. This work was funded by the NCCR Structural Biology (Schweizerischer Nationalfonds).

1. Bissette G, Nemeroff CB, Loosen PT, Prange AJ, Jr., Lipton MA (1976) Hypothermia and intolerance to cold induced by intracisternal administration of the hypothalamic peptide neurotensin. *Nature* 262(5569):607–609.
2. Schimpff RM, et al. (2001) Increased plasma neurotensin concentrations in patients with Parkinson's disease. *J Neurol Neurosurg Psychiatry* 70(6):784–786.
3. Alifano M, et al. (2010) Neurotensin receptor 1 determines the outcome of non-small cell lung cancer. *Clin Cancer Res* 16(17):4401–4410.
4. Griebel G, Holsboer F (2012) Neuropeptide receptor ligands as drugs for psychiatric diseases: The end of the beginning? *Nat Rev Drug Discov* 11(6):462–478.
5. Tanaka K, Masu M, Nakanishi S (1990) Structure and functional expression of the cloned rat neurotensin receptor. *Neuron* 4(6):847–854.
6. Chalon P, et al. (1996) Molecular cloning of a levocabastine-sensitive neurotensin binding site. *FEBS Lett* 386(2-3):91–94.
7. Skrzydelski D, et al. (2003) Differential involvement of intracellular domains of the rat NTS1 neurotensin receptor in coupling to G proteins: A molecular basis for agonist-directed trafficking of receptor stimulus. *Mol Pharmacol* 64(2): 421–429.
8. Savdie C, Ferguson SS, Vincent J, Beaudet A, Stroh T (2006) Cell-type-specific pathways of neurotensin endocytosis. *Cell Tissue Res* 324(1):69–85.
9. Ferguson SS (2001) Evolving concepts in G protein-coupled receptor endocytosis: The role in receptor desensitization and signaling. *Pharmacol Rev* 53(1):1–24.
10. Souazé F, Rostène W, Forgez P (1997) Neurotensin agonist induces differential regulation of neurotensin receptor mRNA. Identification of distinct transcriptional and post-transcriptional mechanisms. *J Biol Chem* 272(15):10087–10094.
11. Toy-Miou-Leong M, Cortes CL, Beaudet A, Rostène W, Forgez P (2004) Receptor trafficking via the perinuclear recycling compartment accompanied by cell division is

necessary for permanent neurotensin cell sensitization and leads to chronic mitogen-activated protein kinase activation. *J Biol Chem* 279(13):12636–12646.
12. Okada T, et al. (2000) X-Ray diffraction analysis of three-dimensional crystals of bovine rhodopsin obtained from mixed micelles. *J Struct Biol* 130(1):73–80.
13. Lebon G, et al. (2011) Agonist-bound adenosine A2A receptor structures reveal common features of GPCR activation. *Nature* 474(7352):521–525.
14. Warne T, et al. (2008) Structure of a beta1-adrenergic G-protein-coupled receptor. *Nature* 454(7203):486–491.
15. Sarkar CA, et al. (2008) Directed evolution of a G protein-coupled receptor for expression, stability, and binding selectivity. *Proc Natl Acad Sci USA* 105(39):14808–14813.
16. Schlinkmann KM, et al. (2012) Critical features for biosynthesis, stability, and functionality of a G protein-coupled receptor uncovered by all-versus-all mutations. *Proc Natl Acad Sci USA* 109(25):9810–9815.
17. Schlinkmann KM, et al. (2012) Maximizing detergent stability and functional expression of a GPCR by exhaustive recombination and evolution. *J Mol Biol* 422(3): 414–428.
18. Scott DJ, Plückthun A (2013) Direct molecular evolution of detergent-stable G protein-coupled receptors using polymer encapsulated cells. *J Mol Biol* 425(3):662–677.
19. Kitabgi P (2006) Functional domains of the subtype 1 neurotensin receptor (NTS1). *Peptides* 27(10):2461–2468.
20. Ballesteros JA, Weinstein H (1992) Analysis and refinement of criteria for predicting the structure and relative orientations of transmembranal helical domains. *Biophys J* 62(1):107–109.
21. Dodevski I, Plückthun A (2011) Evolution of three human GPCRs for higher expression and stability. *J Mol Biol* 408(4):599–615.
22. Choe HW, et al. (2011) Crystal structure of metarhodopsin II. *Nature* 471(7340): 651–655.
23. Rasmussen SG, et al. (2011) Crystal structure of the β2 adrenergic receptor-Gs protein complex. *Nature* 477(7366):549–555.
24. Standfuss J, et al. (2011) The structural basis of agonist-induced activation in constitutively active rhodopsin. *Nature* 471(7340):656–660.
25. Scheerer P, et al. (2008) Crystal structure of opsin in its G-protein-interacting conformation. *Nature* 455(7212):497–502.
26. White JF, et al. (2012) Structure of the agonist-bound neurotensin receptor. *Nature* 490(7421):508–513.
27. Deupi X, Standfuss J (2011) Structural insights into agonist-induced activation of G-protein-coupled receptors. *Curr Opin Struct Biol* 21(4):541–551.
28. Botto JM, et al. (1997) Identification in the rat neurotensin receptor of amino-acid residues critical for the binding of neurotensin. *Brain Res Mol Brain Res* 46(1-2):311–317.
29. Barroso S, et al. (2000) Identification of residues involved in neurotensin binding and modeling of the agonist binding site in neurotensin receptor 1. *J Biol Chem* 275(1): 328–336.
30. Huynh J, Thomas WG, Aguilar MI, Pattenden LK (2009) Role of helix 8 in G protein-coupled receptors based on structure-function studies on the type 1 angiotensin receptor. *Mol Cell Endocrinol* 302(2):118–127.
31. Delos Santos NM, Gardner LA, White SW, Bahouth SW (2006) Characterization of the residues in helix 8 of the human beta1-adrenergic receptor that are involved in coupling the receptor to G proteins. *J Biol Chem* 281(18):12896–12907.
32. Feng GJ, et al. (2003) Selective interactions between helix VIII of the human mu-opioid receptors and the C terminus of periplakin disrupt G protein activation. *J Biol Chem* 278(35):33400–33407.
33. Zhong M, Navratil AM, Clay C, Sanborn BM (2004) Residues in the hydrophilic face of putative helix 8 of oxytocin receptor are important for receptor function. *Biochemistry* 43(12):3490–3498.
34. Kirchberg K, et al. (2011) Conformational dynamics of helix 8 in the GPCR rhodopsin controls arrestin activation in the desensitization process. *Proc Natl Acad Sci USA* 108(46):18690–18695.
35. Lodowski DT, et al. (2007) Crystal packing analysis of Rhodopsin crystals. *J Struct Biol* 158(3):455–462.
36. Faussner A, et al. (2005) The role of helix 8 and of the cytosolic C-termini in the internalization and signal transduction of B(1) and B(2) bradykinin receptors. *FEBS J* 272(1):129–140.
37. Lawrence MC, Colman PM (1993) Shape complementarity at protein/protein interfaces. *J Mol Biol* 234(4):946–950.
38. Wu B, et al. (2010) Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science* 330(6007):1066–1071.
39. Zhang C, et al. (2012) High-resolution crystal structure of human protease-activated receptor 1. *Nature* 492(7429):387–392.
40. Kim TY, Schlieter T, Haase S, Alexiev U (2012) Activation and molecular recognition of the GPCR rhodopsin—Insights from time-resolved fluorescence depolarisation and single molecule experiments. *Eur J Cell Biol* 91(4):300–310.
41. Bruno A, Costantino G, de Fabritiis G, Pastor M, Selent J (2012) Membrane-sensitive conformational states of helix 8 in the metabotropic Glu2 receptor, a class C GPCR. *PLoS ONE* 7(8):e42023.
42. Park JH, Scheerer P, Hofmann KP, Choe HW, Ernst OP (2008) Crystal structure of the ligand-free G-protein-coupled receptor opsin. *Nature* 454(7201):183–187.
43. Mundy DI, Warren G (1992) Mitosis and inhibition of intracellular transport stimulate palmitoylation of a 62-kD protein. *J Cell Biol* 116(1):135–146.
44. Qanbar R, Bouvier M (2003) Role of palmitoylation/depalmitoylation reactions in G-protein-coupled receptor function. *Pharmacol Ther* 97(1):1–33.
45. Chini B, Parenti M (2009) G-protein-coupled receptors, cholesterol and palmitoylation: Facts about fats. *J Mol Endocrinol* 42(5):371–379.
46. Tucker J, Grisshammer R (1996) Purification of a rat neurotensin receptor expressed in Escherichia coli. *Biochem J* 317(Pt 3):891–899.
47. White JF, Trinh LB, Shiloach J, Grisshammer R (2004) Automated large-scale purification of a G protein-coupled receptor for neurotensin. *FEBS Lett* 564(3):289–293.
48. Grisshammer R, White JF, Trinh LB, Shiloach J (2005) Large-scale expression and purification of a G-protein-coupled receptor for structure determination—An overview. *J Struct Funct Genomics* 6(2-3):159–163.
49. Heakal Y, et al. (2011) Neurotensin receptor-1 inducible palmitoylation is required for efficient receptor-mediated mitogenic-signaling within structured membrane microdomains. *Cancer Biol Ther* 12(5):427–435.

3.2. Supporting information

Supporting Information

Egloff et al. 10.1073/pnas.1317903111

SI Text

Materials. The tritiated agonist [^{3}H] neurotensin (NT) ([3,11-tyrosyl-3,5-^{3}H(N)]-pyroGlu-Leu-Tyr-Glu-Asn-Lys-Pro-Arg-Arg-Pro-Tyr-Ile-Leu) and [^{35}S]GTPγS (1250 Ci/mmol) were purchased from Perkin–Elmer. HyLite Fluor 647-labeled NT8-13 (NT8-13-HL647) was purchased from Anawa. Unlabeled NT8-13 (Arg-Arg-Pro-Tyr-Ile-Leu) and NT1 were purchased from Anaspec. N-decyl-β-D-maltopyranoside (DM), 3-[(3-cholamidopropyl)-dimethylammonio]-1-propane sulfonate (CHAPS), n-nonyl-β-D-glucopyranoside (NG), n-decyl-β-D-glucopyranoside (DG), n-dodecyl-β-D-maltopyranoside (DDM), n-octyl-β-D-glucopyranoside (OG), and n-dodecyl-β-D-glucopyranoside (DDG) were obtained from Anatrace. Cholesteryl hemisuccinate tris salt (CHS) and lysozyme were purchased from Sigma. Empty PD10 columns, N-hydroxysuccinimide (NHS)-activated Sepharose, sulphopropyl (SP)-Sepharose, and Superdex 200 10/300 GL were obtained from GE Healthcare. Complete EDTA-free protease inhibitor mixture and DNase I were purchased from Roche. The neurotensin receptor 1 (NTR1) antagonist SR142948 was obtained from Axon Medchem. Sf-900 II serum-free media and Protein G Dynabeads were purchased from Life Technologies. Nickel-nitrilotriacetic acid (Ni-NTA) resin was obtained from Qiagen. Amicon Ultra concentrators, glass fiber filtration plates (MultiScreen-FB), and nitrocellulose filtration plates (MultiScreen$_{HTS}$-HA plates) were purchased from Millipore.

Construct Design. All NTR1 variants were expressed in a pBR322-derived vector, which was constructed from a plasmid originally obtained as a kind gift from R. Grisshammer (National Institute of Neurological Disorders and Stroke, National Institutes of Health, Rockville, MD) (1, 2) and was modified to encode now an N-terminal maltose-binding protein (MBP) linked via a hexahistidine tag and a human rhinovirus (HRV) 3C protease site to G50 (sequential NTR1 numbering). The variants were C-terminally truncated at G390 and linked via a HRV 3C protease site, a penta-asparagine linker, and a di-glycine-serine linker to thioredoxin A (TrxA), which is followed by a deca-histidine tag. Amino acids V280-I295 were deleted in the constructs ΔIC3A and E273-T290 in ΔIC3B. Directed evolution of NTR1 was performed as previously described (3, 4). For mammalian expression, all constructs were subcloned into the vector pcDNA3 (Life Technologies) containing an N-terminal FLAG epitope.

Ligand-Binding Experiments. Whole-cell radioligand-binding assays (RLBAs) in *Escherichia coli* were used to perform affinity measurements. Receptors were expressed in BL21 Tuner cells (Novagen) in 50 mL 2YT medium with 0.3% glucose and 100 μg/mL ampicillin. Cultures were inoculated to OD_{600} = 0.05 and grown at 37 °C to OD_{600} = 0.5. G protein–coupled receptor (GPCR) expression was induced with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) for 20 h at 24 °C or for 16 h at 28 °C, respectively. From the expression cultures, 80 μL (TM86V-ΔIC3A) or 800 μL (NTR1 wild-type) samples were centrifuged, and the pellet was resuspended in 3 mL binding buffer (50 mM Tris·HCl pH 7.4, 0.1% BSA, 1 mM EDTA, and 40 μg/mL bacitracin). We added 20 μL of resuspended cells to 160 μL binding buffer containing [^{3}H]NT at various concentrations for agonist saturation-binding experiments. Antagonist competition experiments were performed using a binding buffer with constant [^{3}H]NT concentrations (10 nM) and various concentrations of the antagonist SR142948.

For determination of expression levels of NTR1 wild-type and TM86V-ΔIC3A in *Spodoptera frugiperda* (*Sf9*) insect cells, membranes were incubated in 200 μL binding buffer containing 10 nM [^{3}H]NT. NTR1 expression in *Sf*9 insect cells and membrane preparation is described below.

For both RLBAs on *E. coli* and *Sf*9 insect cells, all binding reaction mixtures were incubated for 2 h at 4 °C in a 96-well plate on a plate shaker. Note that not all binding reactions for TM86V-ΔIC3A may have completely reached equilibrium after 2 h incubation due to slow on-rates or off-rates. Nonspecific binding was determined in the presence of 10 μM unlabeled NT. The unbound ligand was separated by vacuum filtration on 96-well glass fiber filtration plates (Millipore) that were pretreated with polyethylenimine. Filters were washed four times by 200 μL ice-cold 50 mM Tris·HCl pH 7.4, and they were dried for 1 h at 65 °C. Subsequently, they were dissolved in 200 μL OptiPhase Super-Mix (PerkinElmer) per well. Scintillation was counted on a Microbeta 1450 Plus liquid scintillation counter (Wallac). Data were analyzed by nonlinear regression using GraphPad5 Prism software and fit to one-site binding equations: for agonist saturation-binding experiments after subtraction of background, $y = B_{max} * L/(L + K_d)$; for antagonist competition binding experiments, $y = B + (T - B)/(1 + 10^{-}(x\text{-}\log(IC_{50})))$, where B describes the minimum level of the signal and T the maximum. Less than 10% of free ligand was bound by the receptors in all reactions.

G Protein Purification. G protein composed of $G\alpha_{i1}$, $G\beta_1$, and $G\gamma_1$ was expressed in *Sf*9 cells using a single baculovirus encoding all three subunits, which was generated by using the MultiBac system (5, 6). To purify the heterotrimeric G protein, a 3C-protease–cleavable deca-histidine tag had been fused to the N terminus of $G\beta_1$, whereas an HA tag had been fused to the N terminus of $G\gamma_1$ for immobilization on magnetic beads. *Sf*9 cells grown in Sf-900 II serum-free media were infected at a density of 7×10^6 cells/mL with a multiplicity of infection (MOI) of 5 with the G protein–encoding virus. After 72 h incubation, the infected cells were harvested by centrifugation and resuspended in 30 mL lysis buffer (50 mM Hepes pH 8.0, 50 mM NaCl, 1 mM $MgCl_2$, 10 μM GDP, 5 mM β-mercaptoethanol, complete protease inhibitor mixture) and lysed by sonication. The lysate was centrifuged for 5 min at 500 × *g* to remove cell debris, and the resulting supernatant was ultracentrifuged for 40 min at 108,000 × *g* to collect the membranes. The G protein was purified from the membranes following the procedure described by Rasmussen et al. (7).

Pull-Down Experiments of GPCR by Immobilized G Protein. TM86V-ΔIC3A, as N-terminal MBP and C-terminal TrxA fusion, was expressed in *E. coli* as described above. We incubated 3.1×10^9 cells in PBS containing 1.25 mg/mL lysozyme, 20 μg/mL DNaseI, 1 mM $MgCl_2$, and complete protease inhibitor mixture for 45 min at 4 °C. After the addition of 5 mM EDTA, cells were disrupted by sonication. Cell debris was removed by a low speed centrifugation (10 min at 5,000 × *g*, 4 °C), and cell membranes were pelleted by ultracentrifugation at 118,000 × *g* for 1 h (4 °C). The membrane pellet was resuspended in 1.5 mL 20 mM Hepes pH 8.0, 50 mM NaCl, 30% (vol/vol) glycerol, and flash-frozen for storage at −80 °C.

For the pull-down experiment, 20 μL *E. coli* membranes and 25 μg purified G protein were incubated for 2 h and 40 min, respectively, at 4 °C in 500 μL solubilization buffer [20 mM Hepes pH 8.0, 100 mM NaCl, 2 mM $MgCl_2$, 20 μM NT8-13, 30% (vol/vol) glycerol, 0.6% CHAPS, 0.12% CHS, 1% DDM, complete protease

inhibitor mixture]. Before the binding of the HA-tagged G protein to the anti-HA-antibody–coated (Sigma, H9658) Protein G Dynabeads (5 μg antibody for 900 μg beads/condition), a centrifugation at 20,000 × *g* for 30 min was performed to remove potential protein aggregates. Before incubation of the GPCR with the beads for 1 h, nonsolubilized material was removed by ultracentrifugation at 86,000 × *g* for 30 min. Protein bound to the beads was eluted with standard 1× SDS loading buffer. Between all incubation and elution steps, the beads were washed extensively with solubilization buffer. Eluted proteins were analyzed by Western blot with antibodies against thioredoxin (Sigma, T0803) and $G\alpha_{i1}$ (Lifespan Biosciences, LS-B4007).

Isolation of *Sf*9 Membranes. The GPCRs, tagged N-terminally with tobacco etch virus (TEV) protease-cleavable FLAG and decahistidine tags, were expressed in *Sf*9 cells by infection with GPCR virus at an MOI of 5. At 60 h after infection, cells were lysed by incubation in lysis buffer (10 mM Tris·HCl pH 7.4, 1 mM EDTA, 5 μg/mL Leupeptin, 0.1 mM Pefabloc SC, 1 μg/mL Pepstatin) for 30 min at 4 °C and subsequently forcing them several times through a 27G1/4 needle. After a low-speed centrifugation at 1,000 × *g*, membranes were collected at 20,000 × *g* and incubated for 30 min in buffer A (50 mM Tris·HCl pH 7.4, 1 mM EDTA) containing 7 M urea to remove peripherally bound proteins. The urea concentration was then reduced to 3.5 M by adding buffer A, and the membranes were collected again by centrifugation. The membranes were washed once with buffer A and were flash-frozen for storage at –80 °C in buffer A containing 20% (wt/vol) sucrose. For ligand-binding experiments (i.e., Fig. S4), the urea wash step was omitted.

[^{35}S]GTPγS Assay. The [^{35}S]GTPγS assay was performed with slight changes from a previously described protocol (8). Briefly, urea-washed membranes containing 1 nM of GPCR, which were mixed with 100 nM purified G protein ($\alpha_{i1}\beta_1\gamma_1$) and 200 μM NT or no ligand, were incubated in a total volume of 50 μL of assay buffer (50 mM Tris·HCl pH 7.4, 1 mM EDTA, 100 mM NaCl, 1 mM DTT, 3 mM $MgSO_4$, 0.3% BSA, 2 μM GDP, 4 nM [^{35}S]GTPγS) for 20 min at 25 °C. The reaction was stopped by filtration over nitrocellulose filters (Multiscreen$_{HTS}$-HA plates) and washing four times with ice-cold wash buffer (50 mM Tris·HCl pH 7.4, 1 mM EDTA). Counts arising from G protein alone (non-GPCR–stimulated GTPγS binding) had been subtracted from all reactions. The counts for GPCR-containing urea-washed membranes—that is, without G protein—were all in the same range (Fig. S4). Thus, the higher basal counts (compared with NTR1) observed for TM86V-ΔIC3A and the TM86V-ΔIC3A L167R back-mutant are due to intrinsic properties of the GPCRs, and not due to potentially different membrane preparations. Furthermore, the amounts of membrane used for the mutants had been less than for NTR1.

Large-Scale Expression of NTR1 Variants. A 1 L preculture of *E. coli* BL21 Tuner cells harboring the expression plasmid was grown overnight at 37 °C in 2YT medium, containing 1% glucose and 50 μg/mL ampicillin. A 50 L fermenter (Bioengineering) containing 2YT medium, 0.6% glucose, and 50 μg/mL ampicillin, was inoculated by the complete preculture and grown to an OD_{600} of 2.5 at 37 °C, followed by induction with 1 mM IPTG at 29 °C overnight. Cells were harvested using a continuous-flow centrifuge.

Large-Scale Purification of NTR1 Variants. In a typical purification, 50 g of cell pellet (corresponding to a 7% aliquot of a fermenter run) were resuspended in 100 mL 2× solubilization buffer containing 100 mM Hepes pH 8, 20% (vol/vol) glycerol, 400 mM NaCl, and five tablets of complete EDTA-free protease inhibitor mixture. All subsequent steps were carried out at 4 °C. We added 0.5 mL of 1 M $MgCl_2$, a spatula tip of DNase I, 200 mg lysozyme, 20 mL of CHAPS [6% (wt/vol)]/CHS [1.2% (wt/vol)], and 34 mL of 10% (wt/vol) DM to the resuspended cells while stirring. The mixture was incubated for 15 min. Sonication was performed for 30 min in an ice-water bath. Subsequently, 4 mL of 0.5 M EDTA was added, and the mixture was incubated for another 30 min while stirring. The suspension was centrifuged for 30 min at 15,000 rpm (SLA 1500 rotor), and the supernatant was decanted and mixed to 5 mL slurry pD-NT ligand resin followed by incubation overnight. The mixture was centrifuged at 400 × *g* for 10 min, and the supernatant was decanted. The pelleted resin was packed into an empty PD10 column, and it was washed on a bench-top pressure-flow device with 75 mL of wash buffer 1, containing 25 mM Hepes pH 8, 10% (vol/vol) glycerol, 600 mM NaCl, and 0.5% DM. It was then washed with 40 mL of wash buffer 2, containing 25 mM Hepes pH 7, 10% (vol/vol) glycerol, 150 mM NaCl, 4 mM DTT, and 0.4% NG. The resin was resuspended in a small volume of wash buffer 2, containing 0.7 mg of HRV 3C protease (produced in house), followed by incubation for 2 h. The eluted protein (10 mL) was diluted three times with SP binding buffer, containing 10 mM Hepes pH 7, 10% (vol/vol) glycerol, 4 mM DTT, and 0.3% NG, and it was subjected to another PD10 column containing 5 mL SP Sepharose, which had been preequilibrated with SP binding buffer. The SP resin was washed with 10 mL SP binding buffer, followed by 25 mL SP wash buffer containing 10 mM Hepes pH 7.7, 10% (vol/vol) glycerol, 35 mM NaCl, 4 mM DTT, and 0.3% NG, followed by another 3 mL SP binding buffer. Elution was carried out by ∼15 mL SP elution buffer containing 10 mM Hepes pH 7, 10% (vol/vol) glycerol, 350 mM NaCl, 4 mM DTT, 0.3% NG, and 500 nM NT1. The NTR1 variants were concentrated by an Amicon-15 Ultra concentrator with 50 kDa cutoff to less than 500 μL and loaded on a Superdex 200 10/300 GL column preequilibrated in a buffer containing 10 mM Hepes pH 8.0, 150 mM NaCl, 4 mM DTT, 0.28% NG, and 100 nM NT1. Peak fractions were pooled (1.5 mL) and concentrated by an Amicon-4 Ultra concentrator (50 kDa cutoff) to 3–20 mg/mL.

Vapor Diffusion Crystallization. For TM86V-ΔIC3A, initial crystals were observed at 5 mg/mL protein concentrations in a sitting drop vapor diffusion experiment. The best diffraction quality crystals were obtained, when before crystallization 10% (vol/vol) of X-mix solution 1, containing 10% (wt/vol) NG, 5% (wt/vol) DG, 0.1% DDG, and 1% CHS, was added to the concentrated protein. Crystals were observed after 24 h and grew for about 1 wk in sitting drops containing 1 μL of protein/X-mix solution and 1 μL reservoir solution [50 mM glycine pH 9.4, 1 M NaCl, 26% (vol/vol) PEG 600] at 4 °C. Crystals of TM86V-ΔIC3B and OGG7-ΔIC3A were analogously prepared by using as reservoir a solution of 21.5% (vol/vol) PEG 600, 2 M NaCl, 50 mM glycine pH 9.4, or 22.5% (vol/vol) PEG 600, 0.5 M NaCl, 50 mM glycine pH 9.4, respectively. For HTGH4-ΔIC3A crystallization, the purified protein was mixed with 10% (vol/vol) of X-mix solution 2, containing 10% (wt/vol) OG, 5% (wt/vol) DG, 0.1% DDG, and 1% CHS, and the reservoir solution was 20% (vol/vol) PEG 600, 0.2 M $CaCl_2$, 50 mM NaAcetate pH 5.5. Crystals were harvested by 0.3–0.4 mm loops and cryoprotected by incubation for 3 × 1 min in a solution containing 50 mM glycine pH 9.4, 1 M NaCl, 36% (vol/vol) PEG 600, and 0.3% NG for TM86V-ΔIC3A, TM86V-ΔIC3B, and OGG7-ΔIC3A, or in 36% (vol/vol) PEG 600, 0.2 M $CaCl_2$, 50 mM NaAcetate pH 5.5 for HTGH4-ΔIC3A, respectively. Crystals were flash-frozen in liquid propane.

Data Collection and Structure Determination. All crystals belonged to the space group $P2_12_12_1$, with two protein molecules in the asymmetric unit. Data were collected from a single, cryocooled (100 K) crystal at the beamline ×06SA at the Swiss Light Source with a PILATUS 6 M high-resolution diffractometer and at

the wavelength λ = 1 Å. Data were processed and scaled with the X-Ray Detector Software (XDS) package (9).

The initial structure of TM86V-ΔIC3A was determined by molecular replacement in Phaser (10) using the nociceptin/orphanin FQ receptor structure [Protein Data Bank (PDB) ID code 4EA3] as a search model. The search model was chosen based on the BLAST search of the PDB results (on June 25, 2012) with E-value 4.3087e-19, % Identity = 26.91 and % Positives = 46.48, followed by model preparation in Sculptor (11). Initial maps after molecular replacement were improved by the mr_rosetta protocol (12, 13) with 5 Rosetta (14) models in Phenix (15). Model building was carried out in Coot (16), and refinement was performed with REFMAC5 (17) and phenix.refine (18). The structures of TM86V-ΔIC3B, HTGH4-ΔIC3A, and OGG7-ΔIC3A were determined by molecular replacement using TM86V-ΔIC3A as a search model. The crystallographic statistics (Table S2) were generated with Phenix Graphic Tools (19) and adjusted manually.

Life Cell Microscopy. HEK293T cells were cultured on 18-mm coverslips and transiently transfected with GPCRs and β-arrestin2–YFP. At 48 h after transfection, coverslips were mounted in an imaging chamber and maintained in Tyrode buffer (137 mM NaCl, 5.4 mM KCl, 2 mM $CaCl_2$, 1 mM $MgCl_2$, 10 mM Hepes pH 7.3) at 37 °C. Cells were stimulated with 50 nM NT8–13–HyLite647, and confocal images were taken at indicated time points on a Leica TCS SP5 laser scanning microscope using single-line excitation at 514 nm and 633 nm and an emission bandwidth of 525–580 and 645–720 nm.

1. Tucker J, Grisshammer R (1996) Purification of a rat neurotensin receptor expressed in Escherichia coli. *Biochem J* 317(Pt 3):891–899.
2. Sarkar CA, et al. (2008) Directed evolution of a G protein-coupled receptor for expression, stability, and binding selectivity. *Proc Natl Acad Sci USA* 105(39):14808–14813.
3. Schlinkmann KM, et al. (2012) Maximizing detergent stability and functional expression of a GPCR by exhaustive recombination and evolution. *J Mol Biol* 422(3): 414–428.
4. Scott DJ, Plückthun A (2013) Direct molecular evolution of detergent-stable G protein-coupled receptors using polymer encapsulated cells. *J Mol Biol* 425(3): 662–677.
5. Fitzgerald DJ, et al. (2006) Protein complex expression by using multigene baculoviral vectors. *Nat Methods* 3(12):1021–1032.
6. Bieniossek C, Imasaki T, Takagi Y, Berger I (2012) MultiBac: Expanding the research toolbox for multiprotein complexes. *Trends Biochem Sci* 37(2):49–57.
7. Rasmussen SG, et al. (2011) Crystal structure of the β2 adrenergic receptor-Gs protein complex. *Nature* 477(7366):549–555.
8. Shibata Y, et al. (2009) Thermostabilization of the neurotensin receptor NTS1. *J Mol Biol* 390(2):262–277.
9. Kabsch W (2010) Xds. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):125–132.
10. McCoy AJ, et al. (2007) Phaser crystallographic software. *J Appl Cryst* 40(Pt 4):658–674.
11. Bunkóczi G, Read RJ (2011) Improvement of molecular-replacement models with Sculptor. *Acta Crystallogr D Biol Crystallogr* 67(Pt 4):303–312.
12. DiMaio F, et al. (2011) Improved molecular replacement by density- and energy-guided protein structure optimization. *Nature* 473(7348):540–543.
13. DiMaio F, Tyka MD, Baker ML, Chiu W, Baker D (2009) Refinement of protein structures into low-resolution density maps using rosetta. *J Mol Biol* 392(1):181–190.
14. Leaver-Fay A, et al. (2011) ROSETTA3: An object-oriented software suite for the simulation and design of macromolecules. *Methods Enzymol* 487:545–574.
15. Adams PD, et al. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2): 213–221.
16. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66(Pt 4):486–501.
17. Murshudov GN, et al. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* 67(Pt 4):355–367.
18. Afonine PV, et al. (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr* 68(Pt 4):352–367.
19. Echols N, et al. (2012) Graphical tools for macromolecular crystallography in PHENIX. *J Appl Cryst* 45(Pt 3):581–586.

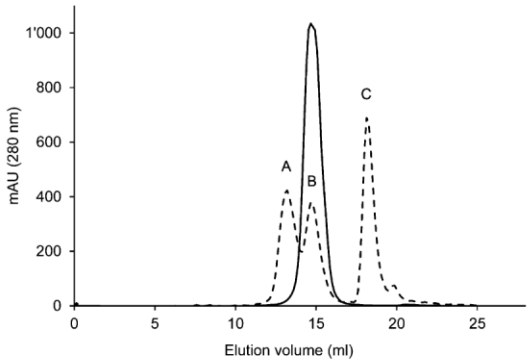


Fig. S1. Preparative size exclusion chromatogram of TM86V-ΔIC3A (solid line) in the short-chain detergent NG. A monodisperse symmetric peak was observed despite the harsh nature of the detergent. The last step of a representative large-scale purification from whole *E. coli* cells with a final yield of about 4 mg purified TM86V-ΔIC3A (from 3.5 L *E. coli* culture) is depicted. Wild-type NTR1 could not be purified under these conditions due to its low stability when solubilized. The harsh short-chain detergents were chosen for crystallization because they are expected to form small micelles around the GPCR. Therefore, they occlude the small hydrophilic domains of GPCRs less than long-chain detergents, and as a result, they allow crystal contact formation. β-amylase (*A*, 200 kDa), BSA (*B*, 66 kDa), and cytochrome c (*C*, 12.7 kDa) were used as reference markers (dashed line). The size exclusion runs were performed by a Superdex 200 10/300 GL column from GE-Healthcare (25 mL column volume).

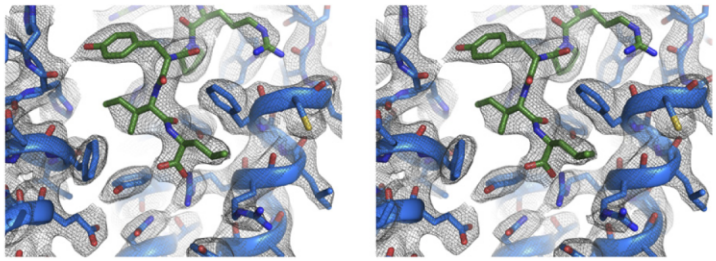
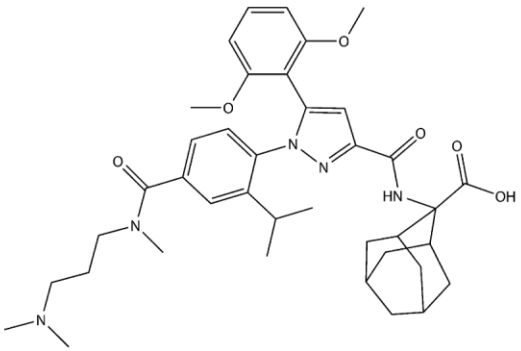


Fig. S2. Quality of the σ_A-weighted $2F_O$-F_C electron density map contoured at 1.2 σ. Stereoview on the five C-terminal neurotensin residues Arg9 to L13 (green) and on parts of transmembrane helix 2 (TM2) (*Left*), TM6 (*Right*), and TM7 (back) of TM86V-ΔIC3B (blue). TM1 and TM3–TM5 (front) were omitted for clarity.

A



B

	TM86V-ΔIC3A	NTR1 wild-type
K_d^{app} NT	2.3 ± 0.4 nM	2.8 ± 0.3 nM
IC_{50}^{app} SR142948	30.0 ± 2.4 nM	8.4 ± 0.9 nM

Fig. S3. Antagonist SR142948 (*A*) and summary table for neurotensin saturation-binding and antagonist competition experiments (*B*). Identical ^{3}H-NT concentrations were applied for wild-type NTR1 and TM86V-ΔIC3A in competition experiments.

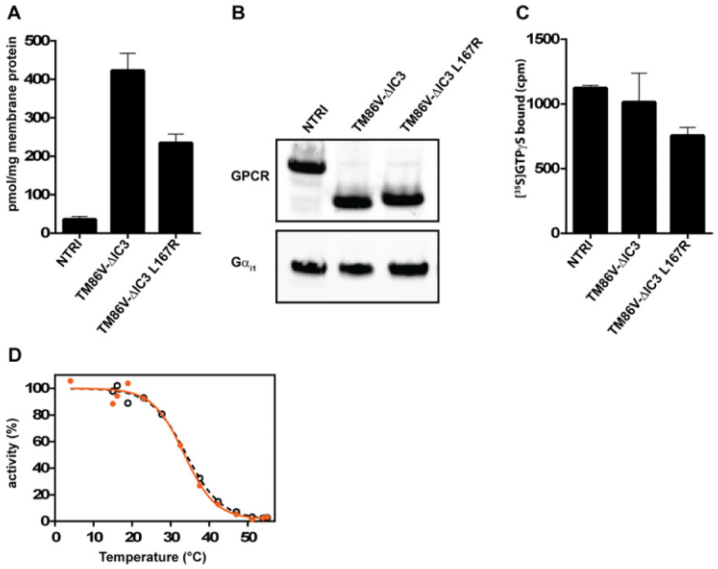


Fig. S4. Functional and total expression levels of wild-type NTR1 and of evolved variants in Sf9 insect cells. (*A*) Radioligand-binding assays with [^{3}H]-neurotensin were carried out to compare the relative functional expression levels of wild-type NTR1, TM86V-ΔIC3A, and TM86V-ΔIC3A L167$^{3.50}$R (back-mutation). As previously reported (1), the higher functional expression levels of the evolved variants in *E. coli* also translate to higher values in eukaryotic expression systems. The back-mutation L167$^{3.50}$R, which restores the highly conserved D/ERY motif in TM86V-ΔIC3A, causes a 40% drop of functional expression. (*B*) Although the functional expression levels of the three protein variants are different, quantitative Western blot experiments of the membranes used in *A* imply that the total expression level is similar. As a reference, Gα_{i1} is shown, which has been coexpressed. (*C*) Background binding of [^{35}S]GTPγS to the membranes only (without added G protein), expressing the different receptors, under same the reaction conditions. Note that the background levels seen here are insignificant (compare cpm here to those in Fig. 2*C*) and that they do not correlate with the amount of membrane applied. This control is important, as for the [^{35}S]GTPγS assay (Fig. 2*A*), we applied equal amounts of functional receptors, and due to the large differences in functional expression levels, different amounts of GPCR-containing membranes had to be used. (*D*) Thermal denaturation profiles of TM86V (orange) and TM86V-L167$^{3.50}$R (black) in the detergent β-D-octyl glucopyranoside. Remaining agonist-binding activity ([^{3}H]-neurotensin) was measured after thermal denaturation in the agonist-bound state for 20 min at the indicated temperature. TM86V displays an apparent T_M of 33.5 °C compared with 34 °C for TM86V-L167$^{3.50}$R. Data of a representative measurement are shown. The method has been described previously (2, 3).

1. Sarkar CA, et al. (2008) Directed evolution of a G protein-coupled receptor for expression, stability, and binding selectivity. *Proc Natl Acad Sci USA* 105(39):14808–14813.
2. Schlinkmann KM, et al. (2012) Maximizing detergent stability and functional expression of a GPCR by exhaustive recombination and evolution. *J Mol Biol* 422(3):414–428.
3. Dodevski I, Plückthun A (2011) Evolution of three human GPCRs for higher expression and stability. *J Mol Biol* 408(4):599–615.

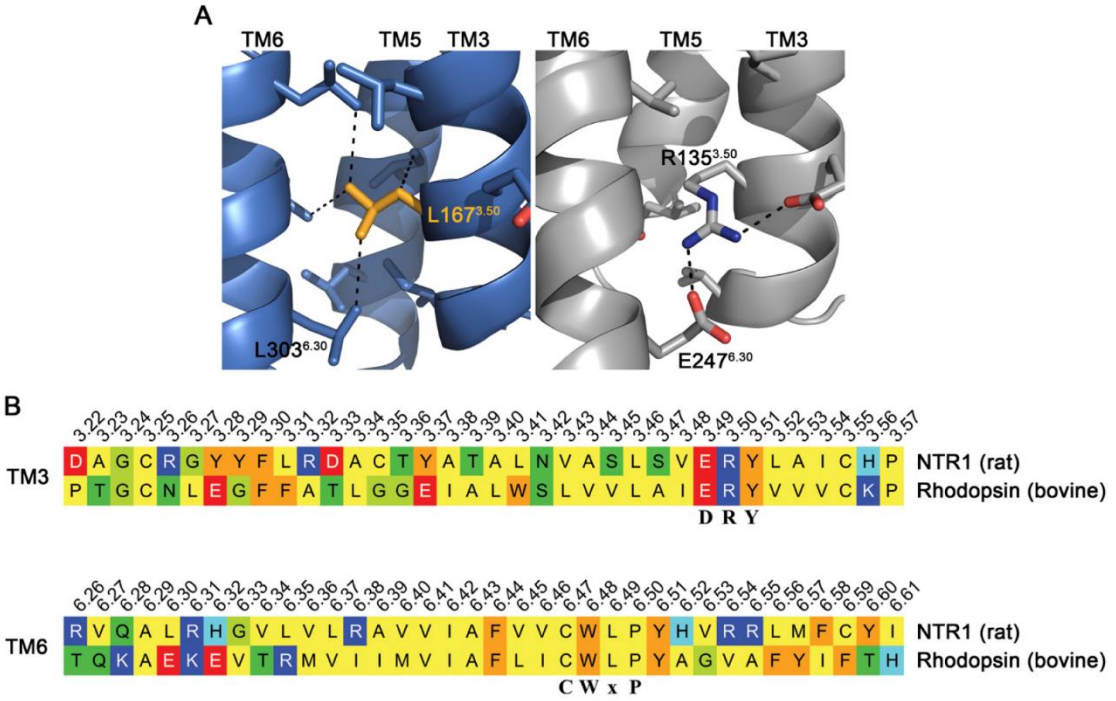


Fig. S5. (*A*) Comparison of the ionic lock of dark-state rhodopsin (*Right*, PDB ID code 1U19) to the arrangement in TM86V-ΔIC3A (*Left*). Wild-type NTR1 does not exhibit an ionic lock, due to the presence of L303$^{6.30}$, and interestingly, directed evolution generated a hydrophobic counterpart by the replacement of the highly conserved R167$^{3.50}$ at helix 3 with leucine (orange). In an evolution-based all-versus-all mutagenesis study, all charged amino acids including the wild-type arginine were found to be counterselected at this position (Fig. S7) (1), and the back-mutation L167$^{3.50}$R decreased functional expression levels also in eukaryotes by about twofold (Fig. S4*A*), thus confirming the positive effect of the selected L167$^{3.50}$ on functional expression levels. Because L167$^{3.50}$ is accommodated in a hydrophobic pocket of TM6 (formed by Leu303$^{6.30}$, V307$^{6.34}$, and L310$^{6.37}$), it appears that the selections favor the naturally occurring "closed" helix arrangement that occludes the putative G protein–binding site (Fig. 4). This implies that conformational flexibility could be a major factor limiting NTR1 expression in *E. coli*. In agreement with this hypothesis, the back-mutation caused strongly increased basal and agonist-stimulated signaling levels toward G_i in TM86V-ΔIC3A (Fig. 2*C*), as it may destabilize the observed inactive state. The increased basal and agonist-stimulated signaling levels of the back-mutant could alternatively be explained by the involvement of the wild-type arginine in direct G_i interactions (2, 3), but as GPCR–G_i complex structures are currently not available, it is unclear whether the selected leucine would directly impair G protein activation. (*B*) Alignment of the rat NTR1 and bovine rhodopsin amino acid sequence of TM3 and TM6. The amino acids were numbered according to Ballesteros–Weinstein (4), and the most conserved residues are indicated. Aliphatic, yellow; aromatic, orange; hydrophilic, green; positively charged, blue; negatively charged, red.

1. Schlinkmann KM, et al. (2012) Critical features for biosynthesis, stability, and functionality of a G protein-coupled receptor uncovered by all-versus-all mutations. *Proc Natl Acad Sci USA* 109(25):9810–9815.
2. Rasmussen SG, et al. (2011) Crystal structure of the β2 adrenergic receptor-Gs protein complex. *Nature* 477(7366):549–555.
3. Scheerer P, et al. (2008) Crystal structure of opsin in its G-protein-interacting conformation. *Nature* 455(7212):497–502.
4. Ballesteros JA, Weinstein H (1992) Analysis and refinement of criteria for predicting the structure and relative orientations of transmembranal helical domains. *Biophys J* 62(1):107–109.

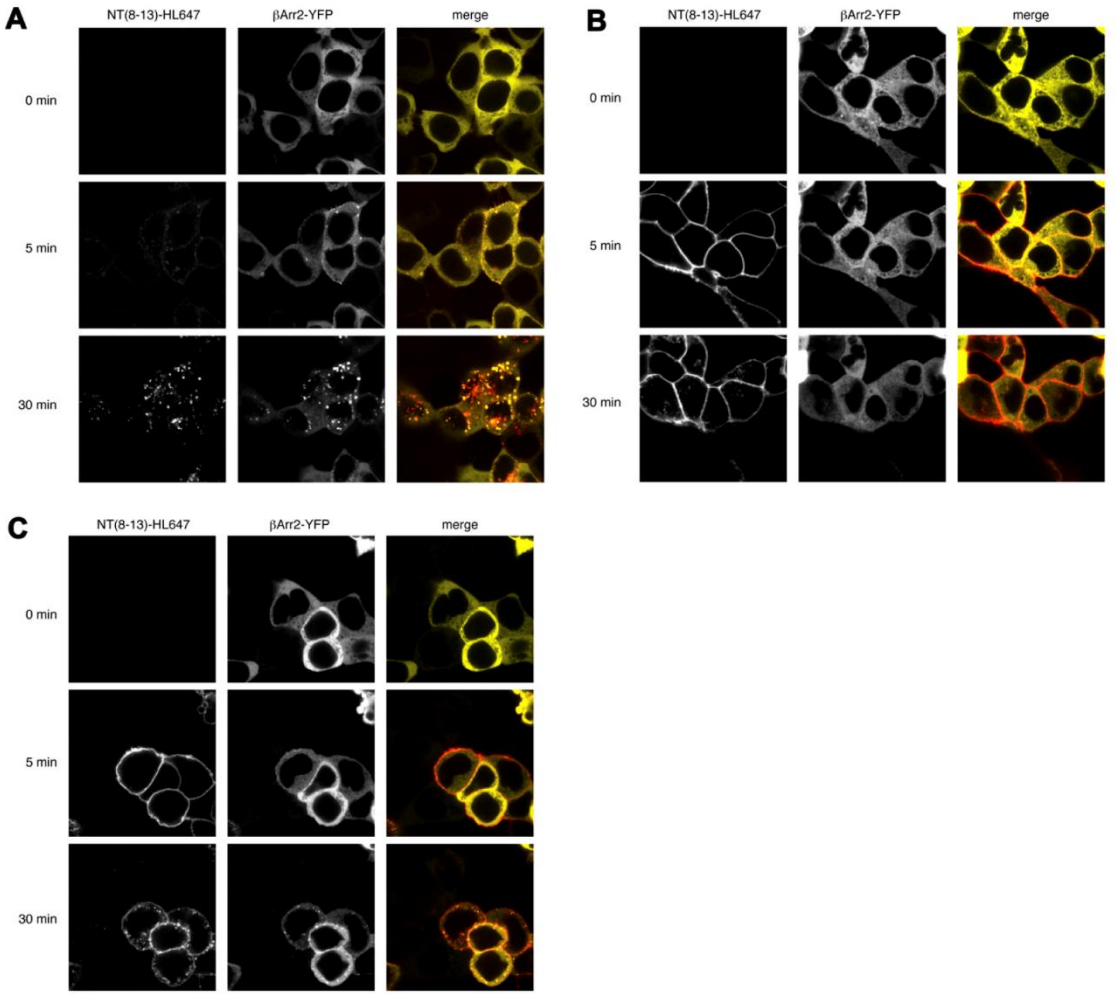


Fig. S6. Agonist binding and β-arrestin2–dependent internalization of NTR1. Confocal images of HEK293T cells coexpressing β-arrestin2–YFP (yellow) with either (*A*) wild-type NTR1, (*B*) TM86V-ΔIC3A, or (*C*) TM86V-ΔIC3A L167R-CT (reconstituted $R167^{3.50}$ and full C-terminal tail), after stimulation with 50 nM fluorescent neurotensin [NT8–13–HL647; red]. All receptor variants bound the agonist readily at the plasma membrane. Wild-type NTR1 strongly cointernalized with β-arrestin2. TM86V-ΔIC3A was internalized weakly, but no clear β-arrestin2 interaction was observed. However, when $Arg167^{3.50}$ of the conserved D/ERY motif as well as the C-terminal tail had been reconstituted (to allow receptor phosphorylation and β-arrestin2 binding), a clear cointernalization was visible. These findings suggest that TM86V-ΔIC3A L167R-CT can interact with β-arrestin2 in a fashion similar to wild type.

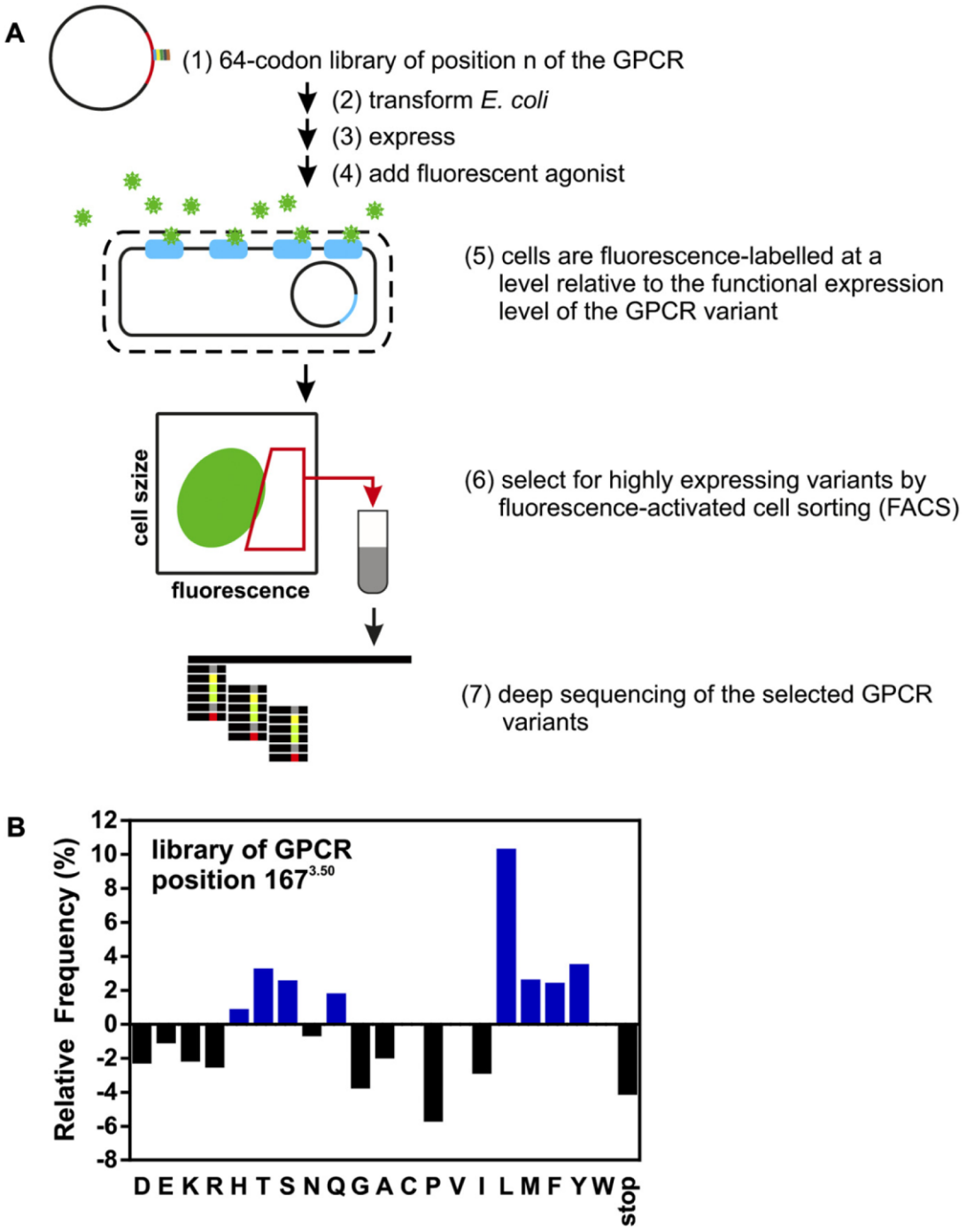


Fig. S7. Functional expression levels investigated by all-versus-all mutagenesis study based on directed evolution. (*A*) Experimental procedure according to Schlinkmann et al. (1): For every amino acid position "n" of the GPCR, a library covering all 64 possible codons at this specific position was transformed and expressed separately in *E. coli* (2, 3). This resulted in a total of 375 individual NTR1 libraries that were processed separately as follows: Fluorescently labeled

Legend continued on following page

neurotensin was used to label the *E. coli* cells relative to their functional expression levels (4, 5). The most highly expressing cells were identified and isolated by fluorescence-activated cell sorting (6). The 375 pools of selected cells were analyzed by 454 sequencing of the GPCR gene segment containing the randomized position (7). We obtained 800,000 sequences in total, which allowed determination of the amino acid preference at every position n. (*B*) 454 sequencing results of the selected library of amino acid position $167^{3.50}$ of NTR1. The difference between the observed amino acid frequency after selection and the input amino acid frequency is plotted. Blue bars, residue types enriched; black bars, residue types counterselected.

1. Schlinkmann KM, et al. (2012) Critical features for biosynthesis, stability, and functionality of a G protein-coupled receptor uncovered by all-versus-all mutations. *Proc Natl Acad Sci USA* 109(25):9810–9815.
2. Sarkar CA, et al. (2008) Directed evolution of a G protein-coupled receptor for expression, stability, and binding selectivity. *Proc Natl Acad Sci USA* 105(39):14808–14813.
3. Schlinkmann KM, et al. (2012) Maximizing detergent stability and functional expression of a GPCR by exhaustive recombination and evolution. *J Mol Biol* 422(3):414–428.
4. Scott DJ, Plückthun A (2013) Direct molecular evolution of detergent-stable G protein-coupled receptors using polymer encapsulated cells. *J Mol Biol* 425(3):662–677.
5. Fitzgerald DJ, et al. (2006) Protein complex expression by using multigene baculoviral vectors. *Nat Methods* 3(12):1021–1032.
6. Bieniossek C, Imasaki T, Takagi Y, Berger I (2012) MultiBac: Expanding the research toolbox for multiprotein complexes. *Trends Biochem Sci* 37(2):49–57.
7. Rasmussen SG, et al. (2011) Crystal structure of the β2 adrenergic receptor-Gs protein complex. *Nature* 477(7366):549–555.

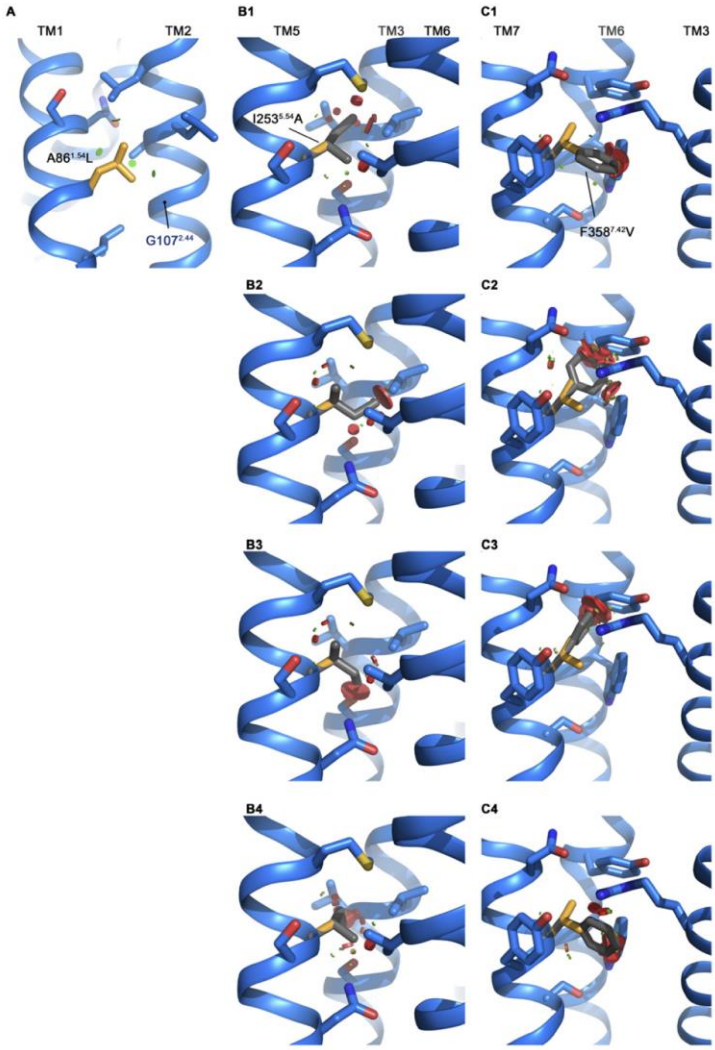


Fig. S8. Close-up view on three thermostabilizing mutations in NTR1-TM86V, completing Fig. 3. (*A*) The selected A86L mutation (orange) allows for additional van der Waals contacts in silico between TM1 and TM2 (green circles). (*B*) The I253A mutation appears to remove clashes (red circles) between residues of TM5, TM3, and TM6. The selected alanine is shown in orange, and the four most common rotamers (1–4) of the wild-type isoleucine are depicted in gray. (*C*) The selected F358V mutation appears to remove clashes between residues of TM5, TM3, and TM6. The selected valine is shown in orange, and the four most common rotamers (1–4) of the wild-type phenylalanine are depicted in gray.

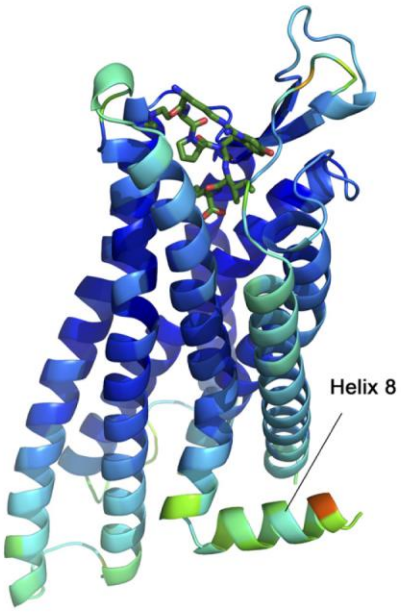


Fig. S9. Relative B-factor coloring on the structure of TM86VΔIC3B. From high to low B-factors: red, yellow, green, and blue. The region with the highest B-factors corresponds to the amphipathic helix 8. Note that NT8-13 is shown independent of B-factors in dark green.

Table S1. Comparison of selected amino acid positions

Sequential	Ballesteros–Weinstein	NTR1-WT	TM86V	HTGH4	OGG7
83	151	S	S	**G**	**G**
86	154	A	**L**	**L**	**L**
101	238	T	T	**R**	**R**
103	240	H	**D**	**D**	**D**
105	242	H	**Y**	**Y**	**Y**
119	256	L	L	**F**	**F**
121	258	M	M	**L**	**L**
124	261	E	E	**D**	**D**
125	262	L	L	L	**V**
143	326	R	R	**K**	**K**
150	333	D	D	**E**	**E**
161	344	A	**V**	**V**	**V**
167	350	R	**L**	**L**	**L**
172	355	C	C	C	**R**
177	360	A	A	A	**H**
208	464	M	M	M	**V**
213	469	R	**L**	**L**	**L**
234	535	V	**L**	**L**	**L**
235	536	K	K	**R**	K
240	541	V	V	**L**	**L**
253	554	I	**A**	**A**	**A**
260	561	I	I	**A**	I
262	563	N	N	**R**	**R**
263	564	K	K	**R**	**R**
305	632	H	**R**	**R**	**R**
313	640	V	V	V	**M**
332	659	C	C	**V**	**V**
342	726	F	F	**A**	**A**
354	738	T	T	**S**	**S**
358	742	F	**V**	**V**	**V**
362	746	S	**A**	**A**	**A**

List of all positions, where mutations occur in at least one of the evolved variants. Bold letters highlight mutations.

Table S2. Data collection and refinement statistics

	TM86V-ΔIC3B	TM86V-ΔIC3A	OGG7-ΔIC3A	HTGH4-ΔIC3A
Data collection				
Wavelength, Å	1.000	1.000	1.000	1.000
Resolution, Å	50–2.75 (2.84–2.75)*	50–3.0 (3.11–3.0)	50–3.1 (3.21–3.1)	50–3.56 (3.7–3.56)
Space group	P 2_1 2_1 2_1	P 2_1 2_1 2_1	P 2_1 2_1 2_1	P 2_1 2_1 2_1
Unit cell				
a, b, c, Å	63.3, 89.4, 212.1	58.6, 90.2, 209.4	60.6, 91.6, 208.6	60.9, 90.9, 211.0
α, β, γ, °	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Multiplicity	13.1 (13.9)	6.9 (5.7)	6.3 (5.6)	7.2 (7.6)
Completeness, %	99.9% (99.8%)	99.8% (98.7%)	99.2% (97.0%)	99.7% (98.4%)
Mean I/σ(I)	19.85 (0.45)	8.68 (0.76)	15.55 (0.51)	8.88 (0.32)
R_{merge}	0.08062 (8.91)	0.1083 (2.821)	0.1026 (3.716)	0.108 (7.401)
$CC_{1/2}$	1 (0.232)	0.997 (0.294)	0.999 (0.126)	0.999 (0.227)
Refinement				
Resolution, Å	19.82–2.75	19.93–3.0	19.88–3.1	21.88–3.57
Total reflections	422,226 (43,292)	158,535 (12,678)	136,489 (11,754)	104,564 (10,617)
R_{work}/R_{free}	0.2478/0.2728	0.2418/0.2792	0.2840/0.3105	0.3092/0.3449
Number of atoms	5,049	4,956	4,865	4,838
Macromolecules	5,049	4,956	4,865	4,838
Ligands	14	5	0	0
Water	0	0	0	0
rms deviations				
Bond lengths, Å	0.003	0.003	0.003	0.003
Bond angles, °	0.65	0.74	0.65	0.78
Ramachandran favored, %	96	99	97	97
Ramachandran outliers, %	0	0	0	0.17
Average B-factor	82	125.7	82.6	189.7
Macromolecules	82	125.7	82.6	189.7
Ligands	91.89	136.28	n/a	n/a
Water	n/a	n/a	n/a	n/a

For each construct the data were collected from a single crystal. Signal to noise of I/σ(I) = 2.0 was at 3.15 Å, 3.26 Å, 3.55 Å, and 4.14 Å for TM86V-ΔIC3B, TM86V-ΔIC3A, OGG7-ΔIC3A, and HTGH4-ΔIC3A, respectively.
*Highest resolution shell is shown in parentheses.

Table S3. Hydrogen bonds and salt bridges between neurotensin ligand and TM86V-ΔIC3A

NTR1 atoms	Distance, Å	Ligand atoms
Hydrogen bonds*		
B:THR 226 [OG1]	2.91	D:TYR 11 [O]
B:TYR 347 [OH]	3.17	D:ILE 12 [O]
B:ARG 327 [NH1]	3.23	D:LEU 13 [O]
B:ARG 327 [NH2]	3.03	D:LEU 13 [O]
B:TYR 146 [OH]	2.70	D:LEU 13 [OXT]
B:GLU 337 [O]	3.40	D:GLY 6 [N*]
B:TRP 339 [O]	3.09	D:GLY 7 [N*]
B:ASP 54 [O]	3.18	D:ARG 8 [NH2]
B:ASP 56 [O]	2.56	D:ARG 8 [NH2]
B:ILE 334 [O]	2.21	D:ARG 9 [NH2]
B:LEU 55 [O]	2.86	D:TYR 11 [OH]
B:HIS 132 [O]	3.25	D:TYR 11 [OH]
Salt bridges*		
B:ARG 327 [NH1]	3.23	D:LEU 13 [O]
B:ARG 327 [NH2]	3.03	D:LEU 13 [O]
B:ASP 336 [OD1]	3.19	D:ARG 9 [NE]
B:ASP 336 [OD1]	3.43	D:ARG 9 [NH1]
B:ASP 336 [OD1]	3.07	D:ARG 9 [NH2]

*Analysis performed using the PDBePISA server. Atoms of linker amino acids of the in-house produced peptide agonist are labeled with an asterisk.

Chapter 4

Discussion and Outlook

Understanding Conformational Dynamics of a Model GPCR – The First Steps

Content

Chapter 4 75
4.1. The potential of crystallized NTR1 variants for GPCR research 76
4.2. NMR 78
4.2.1. Expression in minimal medium and purification considerations 79
4.2.2. Detergent and temperature optimization using ^{1}H- ^{15}N TROSY experiments 80
4.2.1. Evidence of changing receptor dynamics upon binding of different ligands 83
4.3. Outlook and research suggestions 84
4.3.1. Nanodisc reconstitution 84
4.3.1. NMR assignment 86
4.3.2. Generating additional insights into NTR1 function by crystal structures 86
4.3.3. Suggestion: Engineering of mutants with specific signaling characteristics 86
4.4. Materials and Methods 88
4.4.1. Expression in minimal medium (adapted Wagner-lab protocol) 88
4.4.2. NMR experiments on TM86V-ΔIC3 and HTGH4-ΔIC3 89
4.4.3. Circular dichroism (CD) spectroscopy 89
4.4.4. Nanodisc reconstitution 89
4.5. References 91

4.1. The potential of crystallized NTR1 variants for GPCR research

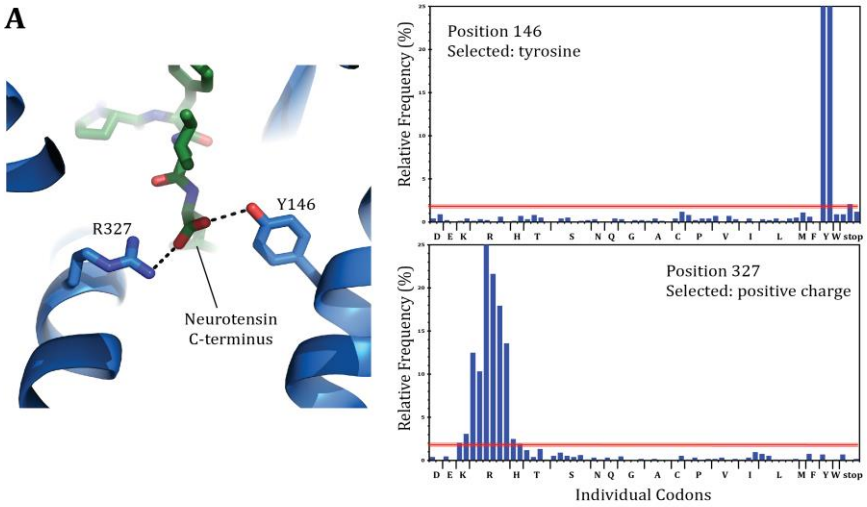
NTR1 is an interesting model for GPCR research, as it signals prototypically via several different G proteins (see section 1.2.1) and since numerous ligands are available, i.e. antagonists and full, partial and biased agonists[1; 2]. Moreover, its natural ligand is a soluble peptide, which can be altered by simple mutagenesis, thus providing an additional handle for experimental manipulations.

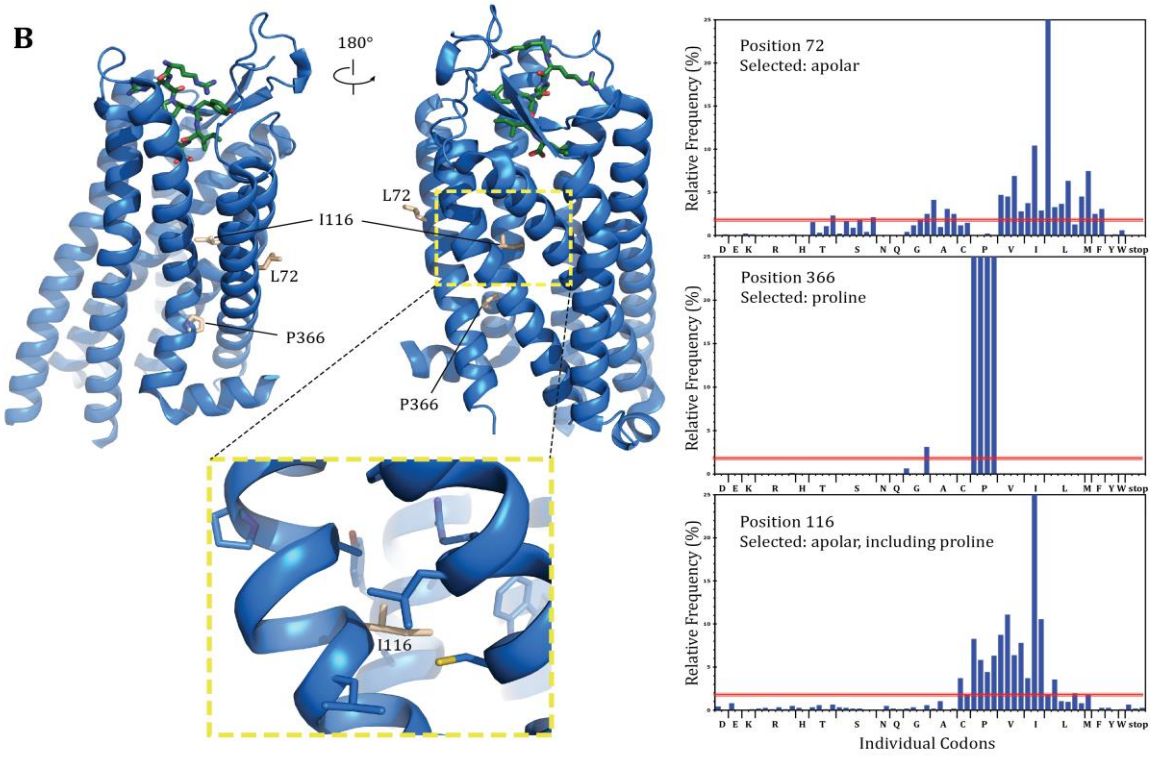
However, like the vast majority of GPCRs, NTR1 is not amenable to research applications that require large quantities of stable proteins in solution, which was the main reason for evolving alternative receptor variants in our research group. The crystal structures reported here provide the first snapshots of a signaling-relevant NTR1 state and they prove that large amounts of highly stable and pure GPCR variants can now be obtained from *E. coli*. The crystallized construct TM86VΔIC3* is of particular interest, as it shares many functional characteristics with wild-type NTR1, like agonist-/antagonist-binding, G_i protein, GRK and β-arrestin interactions, as well as high sequence identity. Given its high stability and expression levels in *E. coli*, it is a unique model receptor that can be used to study general aspects of GPCR function in future experiments *in vitro*.

An additional advantage of TM86V-ΔIC3 is the availability of an extremely large amount of mutagenesis data. Karola Schlinkmann from this research group investigated the effect of all possible single codon substitution at every amino acid position on functional expression levels†, which corresponds to an analysis of approximately 20'000 mutants (also see Chapter 3 Fig. S7)[3]. This body of data is a unique practical guide for the design of future experiments, as it complements structural data in unprecedented ways (examples in Figure 4.1) and furthermore, it inspires many question about fundamental principles of receptor stability, biogenesis, and degradation. NTR1 is to my knowledge the only GPCR, potentially the only protein, where all effects on functional expression levels of *any* single nucleotide or single codon substitution are known.

* As described in chapter 3, two versions of TM86VΔIC3 exist (A and B). Both variants show residual signaling towards G_i (data not shown for TM86VΔIC3B). Additionally, TM86VΔIC3A was tested for internalization, which occurred to some extent. TM86VΔIC3B was not tested in this regard.

† Investigation carried out using the variant D03, which differs at 3 amino acid positions from TM86V.


A
Y146
R327
Neurotensin
C-terminus
Position 146
Selected: tyrosine
Position 327
Selected: positive charge
Relative Frequency (%)
D E K R H T S N Q G A C P V I L M F Y W stop
Individual Codons


B
180°
L72
I116
L72
P366
P366
I116
Position 72
Selected: apolar
Position 366
Selected: proline
Position 116
Selected: apolar, including proline
Relative Frequency (%)
D E K R H T S N Q G A C P V I L M F Y W stop
Individual Codons

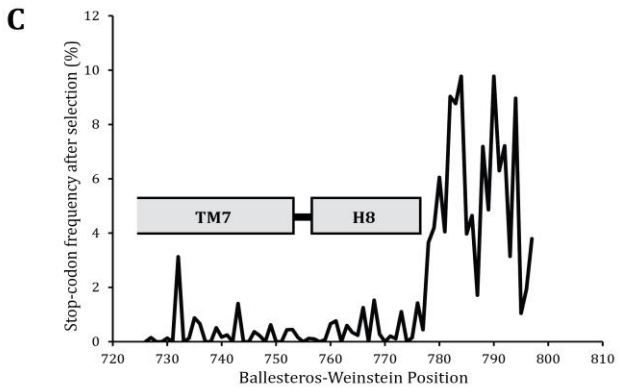

C
Stop-codon frequency after selection (%)
TM7
H8
Ballesteros-Weinstein Position

Figure 4.1 Examples of available sequence statistics in the structural context. The relative frequency of each codon after one round of FACS is displayed by blue bars in the diagrams (right). The red line depicts the expected frequency in the absence of selection pressure. High bars imply favorable codons for elevated functional expression levels at the indicated position and vice versa (see Chapter 3 Fig. S7 and Schlinkmann et al. (2012)[3] for details). Note that the codon with the highest bar is in each plot an expected artifact (wild-type codon), which is due to the pooling of libraries – it is not to be taken into account. A, Typical examples from the ligand-binding pocket, illustrating the preservation of the neurotensin-contacting residues upon directed evolution. B, Examples from the transmembrane region. The side chain of the amino acid at position $72^{1.40}$ points into the membrane space and it is thus selected to be an apolar residue. Proline $366^{7.50}$ is located far away from the ligand-binding site and hence from the selection pressure. Nevertheless, it is the only acceptable amino acid for functional expression at this position, suggesting that its helix-destabilizing properties are of crucial structural relevance. Position $116^{2.53}$ is an example for an unexpected result in terms of amino acids selections. Here, wild-type NTR1 harbors an isoleucine, but interestingly, a proline residue would also be allowed in spite of its position in the middle of the α-helix. The toleration of a non-wild-type proline is a unique exception among all positions. C, Stop-codons are counter-selected throughout the GPCR, but their relative frequencies increase directly after helix 8. Hence, the presence of the amino acid stretch corresponding to Helix 8 is promoting functional expression levels, which might be due to its relevance for the structural integrity of the entire GPCR.

4.2. NMR

For the β_2-adrenergic and many other receptors, pronounced ligand-dependent structural changes are not observed by x-ray crystallography, implying that receptor dynamics are an important key to the understanding of signaling mechanisms. Similarly, the structures of TM86V-ΔIC3 reported here are at the extracellular side in an active, agonist-bound state, but intracellularly, they are clearly exhibiting a prototypic inactive conformation. The crystallized constructs can nevertheless activate G_i proteins to some extent, which is consistent with a high degree of conformational flexibility or the sampling of multiple inactive or active states that was previously reported for other receptors[4].

A number of biophysical studies on dynamics and conformational exchange were reported in the case of rhodopsin and for the β_2-adrenergic receptor, including NMR and DEER (Double electron-electron resonance) spectroscopy[4-11]. Many of these studies provided strong and direct evidence for the existence of alternative TM6 conformations in detergent micelles, which was interpreted in the context of crystal structures. One limitation of these approaches was the use of chemical probes (e.g. spin-labels for DEER or fluorine / ^{13}C-methyl labels for NMR) at specific sites that allowed to monitor changes only at very few engineered spots in the receptor.

In principle, NMR has, however, the potential to follow alterations in chemical environments more globally, if appropriate labeling schemes are applied and given that the spectra are of sufficient quality for complete assignments. The latter is particularly challenging in the case of GPCRs: large proteins, such as MPs in micelles, exhibit comparably slow molecular

tumbling, which causes short transverse relaxation times T_2 (reduced spin coherence lifetimes) and thus low sensitivity and spectral line broadening[12]. In order to prevent rapid magnetization decays to some extent, soluble proteins are usually perdeuterated (> 98 %) and spectra are recorded at elevated temperatures[13-17]. Both of these strategies are inapplicable in the case of the vast majority of GPCRs, as they are unstable and because they cannot be overexpressed in *E. coli* – one of the very few organisms, which would allow full deuteration of proteins.

As the crystallized NTR1 variants fulfill both of these criteria, they represent promising candidates, where complete assignments might be obtainable. This would be a starting point for a series of experiments that were not possible on GPCRs so far: First, local changes of chemical environments may be observed throughout the entire GPCR upon binding of different ligands. Second, it may allow to map the binding sites of further ligands or even to determine ligand-bound structures, if sufficient intra- and intermolecular NOEs can be observed. Third, it may allow to map the binding sites of several different G protein variants and to determine relevant structural aspects for downstream receptor specificity. Analogous experiments might be possible for GRKs and β-arrestins. And fourth, new insights might be generated on conformational dynamics, rate constants can potentially be deduced and the population of specific states may be estimated under different conditions. This chapter describes initial experiments that were conducted with the aim of assigning backbone resonances of evolved NTR1 variants.

4.2.1. Expression in minimal medium and purification considerations

To obtain an initial estimate of expression yields in minimal medium for TM86V-ΔIC3B, a number of expression conditions in shake flasks were tested in the Plückthun laboratory using unlabeled minimal medium and the functional expression levels were assayed by radioligand-binding experiments (work of Jendrik Schoeppe). The main conclusion after an array of optimizations was that the final OD was generally lower than in rich medium (as expected), but the functional expression levels per cell were usually higher. The total functional expression per liter of culture was lower than in rich medium (approximately 1 mg/L for TM86V-ΔIC3B), but nevertheless, it was sufficient to tackle the next steps.

A concern for most biophysical studies including NMR is the functional state of the purified protein. A significant fraction of expressed protein may be non-functional, due to denaturation during purification or because of improper folding during biosynthesis[5]. The latter may be particularly problematic for evolved variants, as the selection pressure was on maximal functional expression levels and not on minimal misfolded receptors. In other words, if a gain in functional expression is due to higher *total* expression levels of a particular variant, as documented for the parent mutant D03[18], it is possible that the amount of misfolded receptors

per cell is also increased. Considering this, it is worth mentioning that standard IMAC purifications sometimes resulted in relatively monodisperse peaks in subsequent SEC analyses (data not shown) for the most stable NTR1 variants in DDM (not for D03), which is certainly an indication for a very low aggregation tendency in this particular detergent. Due to the small protein size compared to the large micelles, SEC gives however very little information on the structural integrity of non-aggregated GPCRs, implying that non-native – but nevertheless soluble – receptors can well go undetected. In crystallization trials, this type of sample inhomogeneity is disadvantageous, but it does not necessarily prevent crystal genesis entirely, as lattice formation itself can act as a purification step selecting for natively folded receptors. For NMR, on the other hand, sample inhomogeneity inevitably reduces spectral quality, as the entire protein ensemble of a sample is under investigation. In my opinion, it is thus very likely that ligand column-based purifications, which isolate functional proteins only, are beneficial for biophysical experiments.

In order to enable our collaborators at Harvard Medical School (Wagner group) to perform optimal NMR experiments, the ligand column technology (see Chapter 2) was transferred to their laboratories. The transfer included the pD-NT construct, receptor plasmids (TM86V-ΔIC3B, HTGH4-ΔIC3B), various protocols and hands-on practical teamwork in Boston together with Dr. Franz Hagn, who is working independently on this project by now. Note that all NMR experiments described herein were performed by Dr. Hagn with me as a purification specialist accompanying many measurements.

4.2.2. Detergent and temperature optimization using ^{1}H- ^{15}N TROSY experiments

NMR spectral qualities are critically affected by the size of the receptor-micelle complex and by the stability of the receptor. There is a potential trade-off between the use of short-chain detergents (small micelles) that are rather destabilizing, and the application of amphiphiles with longer aliphatic tails (large micelles), which are milder. Moreover, alternative detergents can affect NMR data quality also in other ways: They may influence conformational GPCR dynamics and varying exchange rates at the detergent-protein interface can be exhibited, so that specific resonances may either be visible or not. It is non-obvious, which one of the parameters is the most critical for NMR spectral quality, implying that empirical testing of alternative detergents is required to optimize the spectra for assignment.

In order to collect the first NMR spectra in the Wagner laboratories and to determine whether the project is feasible, the most stable NTR1 variant was used (HTGH4). HTGH4-ΔIC3 was expressed in ^{15}N-labeled M9 minimal medium using D_2O as a solvent. The receptor was purified according to the principles described in Chapter 2, but with detergent exchange at the SP sepharose column stage. Milligram quantities of perdeuterated ^{15}N-labeled HTGH4 were

prepared in either 0.05 % (w/v) DDM, 0.3 % (w/v) DM, 0.2 % (w/v) OGNG and 0.1 % (w/v) DH_7PC. Small fractions of the purified receptors were used to record irreversible temperature unfolding curves using circular dichroism (CD) spectroscopy (Figure 4.2). As expected, the maltoside detergents allowed high thermostabilities, but they generated comparably large micelles. OGNG was previously observed to be rather mild in crystallization trials, while still generating small micelles. This could indeed be confirmed here, as HTGH4-ΔIC3 exhibited the smallest apparent molecular weight in the SEC analysis in OGNG and at the same time it was of similar thermostability as in DM. DH_7PC is frequently used as a membrane mimic for NMR experiments. This amphiphile resulted in slightly smaller micelles compared to DM, but the receptor unfolded at 3°C higher temperature.

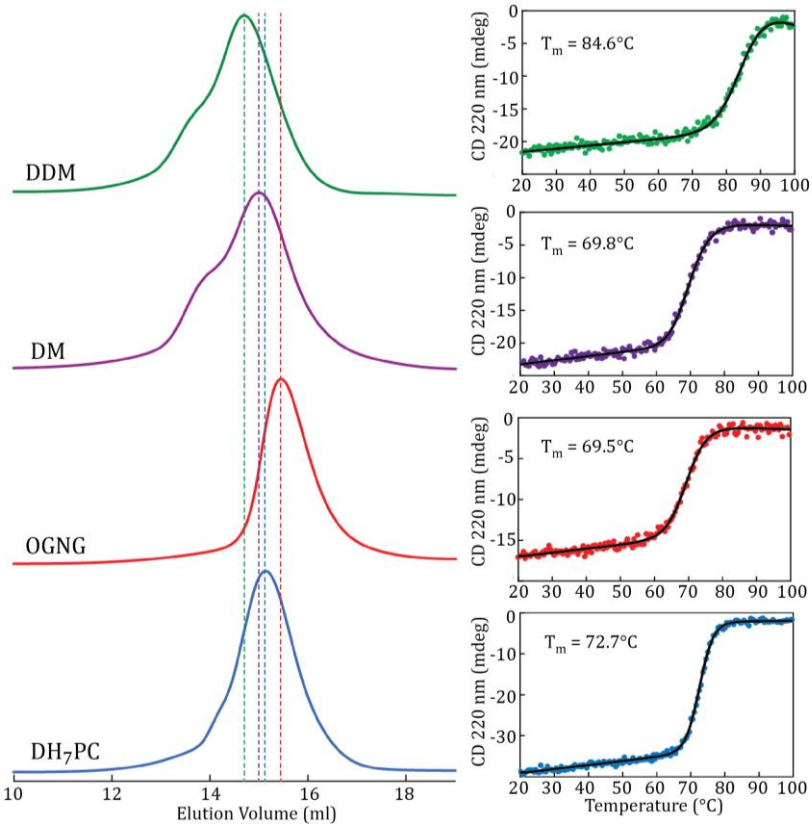


Figure 4.2 Purification and thermostability analyses of HTGH4-ΔIC3 in alternative detergents. Quantitative S200 analyses (left) representing the final step of a purification procedure according to chapter 2. Dimerization shoulders are observed for the maltoside detergents and to a lesser extent for DH_7PC. The relative shoulder heights are concentration dependent, but reducing agent independent (data not shown). Oligomerization is potentially disadvantageous for NMR, due to the expected decrease in molecular tumbling rates and further loss of sensitivity. No shoulders are observed in OGNG. The thermal unfolding curves are based on helical content measurements by CD spectroscopy (right). The cooperativity of the irreversible transitions strongly implies that helical content is a valid relative measure of structural integrity in the case of HTGH4-ΔIC3. DH_7PC appears to be a reasonable compromise between particle size and thermostability. This figure was adapted from a report prepared by F. Hagn.

^{1}H-^{15}N TROSY spectra were recorded in the neurotensin-bound states in the four detergent conditions. Generally, the spectral qualities were highly promising, as a large fraction of the expected amide resonances could be observed. As expected from the unfolding transitions above 69°C, the spectral qualities improved for all detergent conditions with increasing temperatures up to 47°C (Figure 4.3). Surprisingly, however, the detergent exhibiting the smallest micelles (OGNG) produced the poorest spectrum, which might have been due to adverse receptor dynamics in this detergent or because of unfavorable exchange kinetics of this particular amphiphile with the receptor. Comparing the spectra in both maltosides, it appeared that the peak numbers were very similar, but the resonances were significantly better defined in DM. This implied, that the micelle size was the limiting parameter regarding spectral quality in the case of DDM and not receptor stability, which was in agreement with the thermostability analysis. The most useful behavior of HTGH4-ΔIC3 for NMR assignments appeared to occur in DH_7PC, as the largest fraction of expected backbone amide resonances (approximately 250 out of 330) was visible in these TROSY spectra (e.g. compare the number of amide resonances from glycines above 8 ppm (δ_2 ^{1}H) and below 113 ppm (δ_1 ^{15}N)) and since the peaks were well defined. A qualitative comparison of the HTGH4-ΔIC3 spectrum in DH_7PC at 47°C with a previously assigned spectrum of a protein of similar size ($G\alpha_i$, data not shown), suggested that it might be feasible to assign the majority of HTGH4-ΔIC3 resonances to specific backbone amides in this condition using triple labeling strategies and higher dimensional NMR experiments.

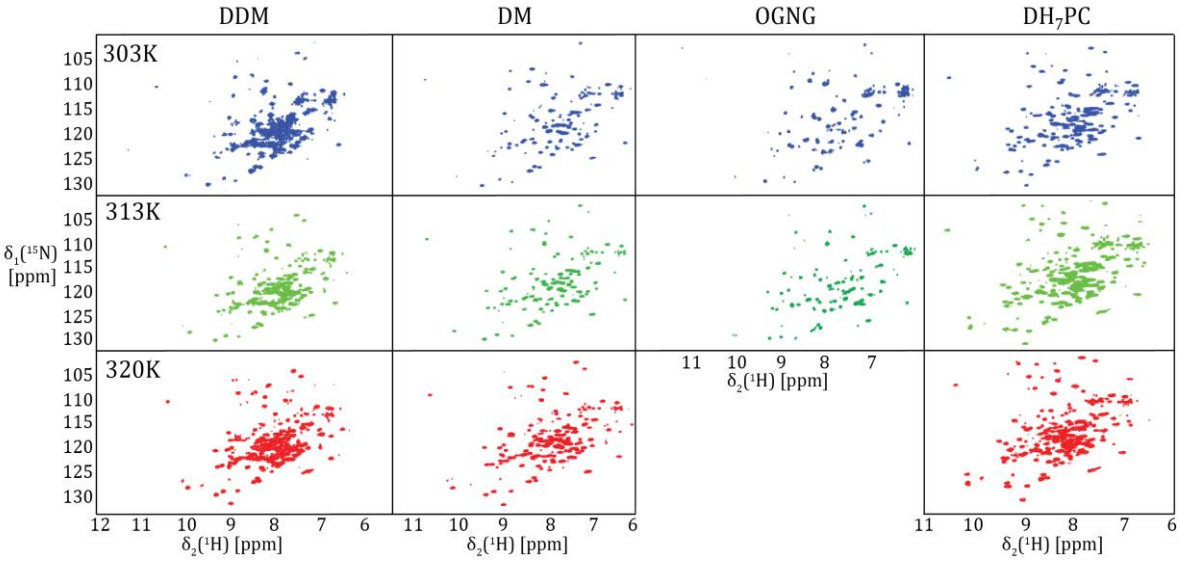


Figure 4.3 ^{1}H-^{15}N TROSY spectra of HTGH4-ΔIC3 at different temperatures and in different detergents. The receptor-neurotensin complex concentrations were at 100 to 200 µM and the spectra were recorded over night. The same sample was used for all three spectra of one detergent condition. Experiments were performed in a sequence from low to high temperatures. Approximately 250 resonances are observable (maximally expected ≈330). This figure was adapted from a report prepared by F. Hagn.

4.2.1. Evidence of changing receptor dynamics upon binding of different ligands

One of the first experiments that was carried out after condition optimization was antagonist competition. As described above, the initial ^{1}H-^{15}N TROSY spectra were obtained after applying the best established ligand-purification method, which resulted in preparations of neurotensin-bound receptors (HTGH4-ΔIC3 in complex with the unlabeled cleavage product from the ligand column). Subsequently, 2 mM unlabeled NTR1-antagonist (SR142948) was added to the NMR sample (100 µM HTGH4-ΔIC3 in DH_7PC) and another ^{1}H-^{15}N TROSY spectrum was recorded (Figure 4.4B). Interestingly, a large fraction of peaks either disappeared or changed their position. The fact that not all resonances were perturbed, implied that HTGH4-ΔIC3 might still have retained a native fold after adding excess antagonist. Additionally, Dr. Hagn could show that antagonist binding leads to an almost 10°C increase of the thermal unfolding transition (Figure 4.4A), which is in agreement with native receptor-antagonist interactions as well. If the ^{1}H-^{15}N TROSY spectrum in the presence of excess antagonist indeed represents an antagonist-bound and natively folded receptor, the missing resonances are likely due to an increased fraction of backbone amides in intermediate exchange. Few missing peaks may be explained by direct effects at the ligand binding site, but the large extent of peak disappearance suggests that the agonist-antagonist replacement must have indirect effects on major parts of the receptor and its dynamics, including residues that are clearly remote from the ligand binding sites.

Notably, in the Plückthun laboratory, an array of experiments suggested that neurotensin binds practically irreversibly or with very low off-rates to the evolved variants (examples in chapter 2). This is not necessarily contradicting the NMR ligand-competition experiment, as the assays in the Plückthun group were performed at 4 – 20°C in maltoside or glucoside detergents for crystallographic purposes. The NMR and CD spectroscopy experiments in the Wagner laboratories, were, however, carried out using receptors prepared in DH_7PC and at much higher temperatures, which may allow increased neurotensin off-rates. In fact, a similar competition experiment in DDM did not reveal particular spectral differences (personal communication F. Hagn), suggesting that the choice of detergent indeed critically affects the off-rate of neurotensin.

A technical challenge of NMR experiments in the context of membrane proteins should, however, be considered as well. Apparently, chemical shifts and their perturbations are dependent on the environment of the nuclei under investigation. When analyzing the effect of a ligand, it is crucial that the ligand and the protein can be provided in exactly the same solvent, otherwise global or potentially also local perturbations can occur, which are ligand-independent. This is challenging for MPs, when detergents are used as membrane mimics, since the protein

concentration step usually enriches empty micelles in an unpredictable way. Without determining the total detergent concentration in the MP sample by a separate assay, it is thus hardly possible to dissolve the ligand in exactly the same buffer and large differences in detergent concentrations may occur. In my opinion, it is therefore important to verify the results of the antagonist competition experiment using an alternative membrane mimic, which does not cause this particular problem.

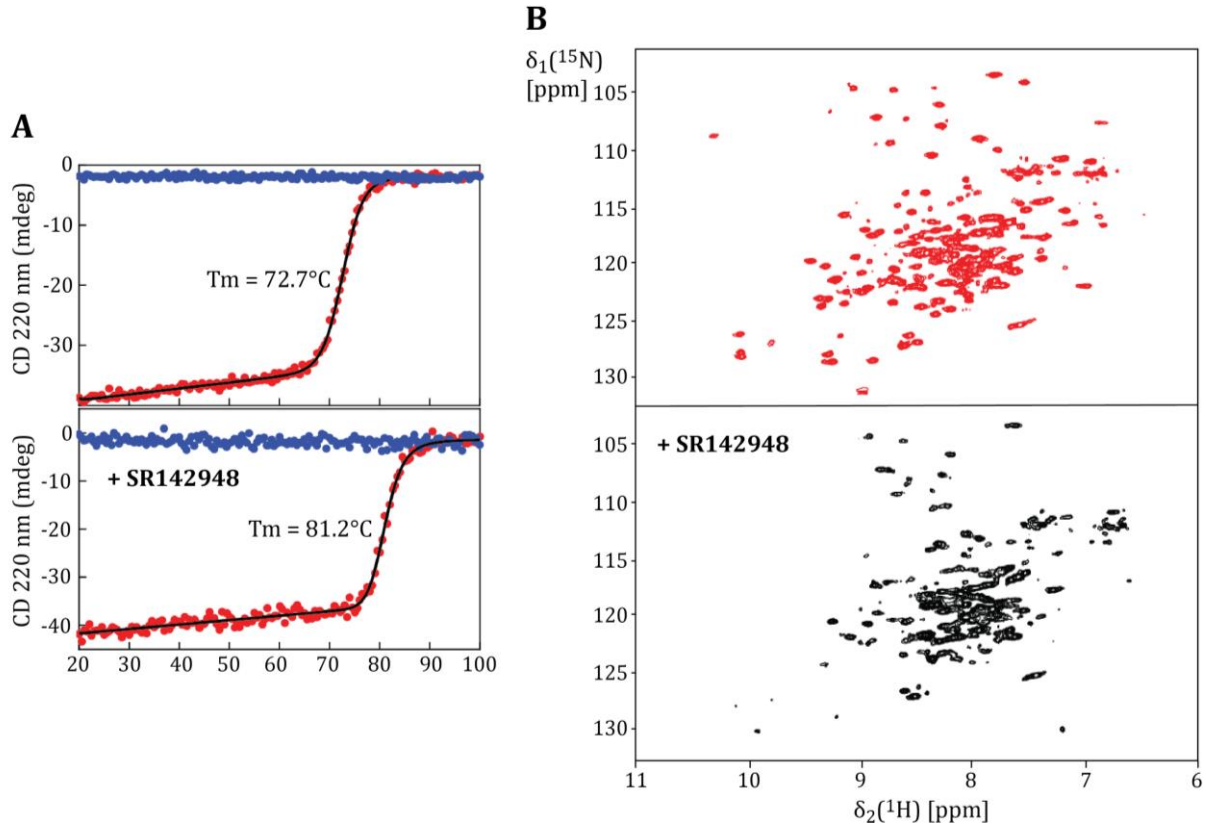


Figure 4.4 Effect of antagonist SR142948 on thermostability and on the ^{1}H-^{15}N TROSY spectrum of HTGH4-ΔIC3 in DH_7PC. These experiments were carried out by Dr. F. Hagn based on samples that were purified according to the principles described in Chapter 2, which resulted in neurotensin-bound receptor preparations. A, the addition of excess antagonist leads to an increased thermal denaturation transition point. As only one transition is observed, it can be concluded that all receptors are bound to antagonist under these condtions. If only a fraction of the receptors would be in complex with the antagonist, two transitions would be expected. B, The spectrum of HTGH4-ΔIC3 at 100 µM in complex with unlabeled neurotensin before (top) and after adding 2 mM unlabeled antagonist (bottom). This figure was adapted from a report prepared by F. Hagn.

4.3. Outlook and research suggestions

4.3.1. Nanodisc reconstitution

A nanodisc is a patch of phospholipids of defined size, which is encircled by two copies of a membrane scaffold protein (MSP)[19]. It is possible to reconstitute MPs into this bilayer particle, which allows handling of the proteins in the absence of detergents[20]. A membrane mimic of this

type could be advantageous for NMR titration experiments, as it is a stable particle that – unlike detergent micelles – is not constantly exchanging components with the surrounding solution at high rates. Hence, empty nanodiscs can be removed after reconstitution, if appropriate chromatographic steps can be identified. Subsequently the "occupied" nanodiscs can be treated like soluble proteins, which is advantageous for NMR titrations, as well as for many other applications, where detergents may interfere (e.g. ligand screening, functional assays, ribosome display, etc.).

During my stay in the Wagner laboratory in Boston, we started to develop reconstitution protocols based on previously described procedures and based on the practical experience of the lab members[21; 22]. After optimizing ratios between lipids, MSP and receptors, we identified conditions that allowed efficient reconstitution of evolved receptors into nanodiscs in small-scale. Furthermore, a highly effective cation-exchange chromatography step was developed, which allowed separation of empty nanodiscs from those containing the receptors (Figure 4.5 . As analyzed by SEC, the small-scale nanodisc reconstitution resulted in homogeneous particles (Figure 4.5 A). In order to apply the nanodisc technology to the various studies in the near future, it is now necessary to establish scaled-up procedures that allow efficient reconstitution of mg quantities of evolved receptors.

Nanodisc reconstitution is a multiparameter reaction that depends on a variety of practical aspects. According to my experience, the procedure described in the methods (section 4.4.4) is highly reproducible as long as the same lipid and MSP stock solutions can be used. If new solutions are prepared (which was necessary during the optimizations in the Wagner laboratory) unwanted variability can occur. In order to establish the large-scale reconstitution reaction in the Plückthun group, I therefore recommend to run initially some pretests at small-scale using similar MSP to lipid ratios to the ones we established in the Wagner laboratory. When the optimal MSP to lipid ratios are identified with the new stock solutions, large-scale reactions based on the same stocks can be established.

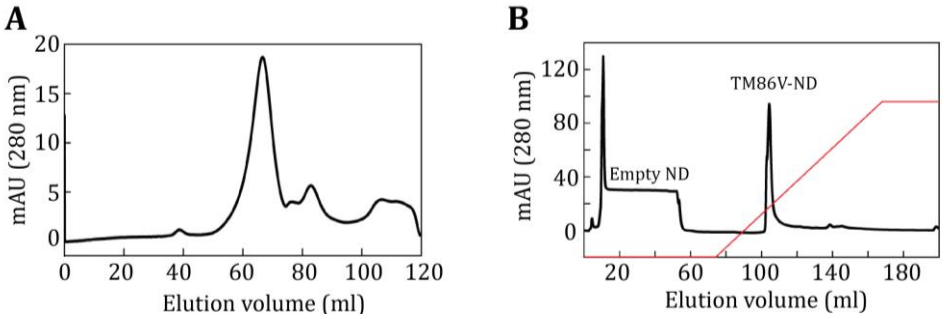


Figure 4.5 Small-scale reconstitution of HTGH4-ΔIC3 into nanodiscs. A, S200 (HiLoad 16/600) analysis directly after reconstitution (detergent removal by dialysis). The main peak corresponds to empty nanodiscs and receptor-containing nanodiscs. The peak at an elution volume of approximately 80 ml is due to the desired excess of MSP in the reconstitution reaction. The void peak at 40 ml and potential oligomerization shoulders are minor. B, Subsequent, cation exchange step, which allows specific binding of nanodiscs that contain receptors and their elution in a small volume. Due to the low pI of MSP, empty nanodisc cannot bind to the column material at the chosen pH (verified by SDS-PAGE analysis).

4.3.1. NMR assignment

The most promising ^{1}H-^{15}N TROSY spectra were obtained using the variant HTGH4-ΔIC3B. TM86V, the only mutant that was characterized in terms of G_i signaling and β-arrestin-mediated internalization, was tested as well (construct TM86V-ΔIC3B). The obtained spectra of the latter were very similar, but of slightly lower quality (data not shown). The two variants are only different at 12 amino acid positions, suggesting that the assignment should be carried out using HTGH4-ΔIC3. It may then be transferred to the signaling-active TM86V-ΔIC3B in a subsequent step. Alternatively, HTGH4-ΔIC3B, which was not characterized in signaling assays so far, can be characterized in terms of function. If experiments reveal that this variant exhibits residual signaling activity as well, a transfer of the assignment may not be necessary.

4.3.2. Generating additional insights into NTR1 function by crystal structures

To improve our understanding of receptor function, it is crucial to obtain further high-resolution structures of NTR1 mutants in alternative states. Mattia de Luigi and Philipp Heine from the Plückthun group are currently working on the crystallization of antagonist-bound TM86V-ΔIC3B and HTGH4-ΔIC3B. Using the mild LCP method, it may even be possible to obtain apo-state structures. Moreover, Jendrik Schoeppe, Abinav Kumar and Matthias Hillenbrand are establishing large-scale complex formation procedures of NTR1 variants with heterotrimeric G proteins. The SEC chromatograms of their complexes look promising as well, but further work is required to obtain crystals.

4.3.3. Suggestion: Engineering of mutants with specific signaling characteristics

The extensive protein engineering efforts for crystallographic studies on GPCRs in the past years included the introduction of fusion proteins, mutations and artificial disulfide bridges.

Interestingly, for those GPCRs, where several approaches were successful, major global effects of these modifications on the structures were rarely observed. But at the same time, signaling activity was usually affected very strongly. TM86V-ΔIC3 is one of the rare cases that exhibits residual signaling towards G proteins, and considering its benefits in terms of expression and stability, it could be a starting point to reengineer *specific* signaling characteristics of the wild-type receptor. The motivation behind this suggestion is explained in the following paragraphs.

Over the course of natural evolution, wild-type NTR1 experienced likely only a very weak selection pressure for high stability. It evolved to couple to various different G proteins and to interact with GRKs, β-arrestins and potentially with other regulatory proteins. Binding to various signaling partners likely requires structural plasticity, which could be one of the reasons for the observed instability of NTR1 (even in the mildest detergents). Low stability may even be a functional requirement that allows rapid receptor turnover upon activation, which is typical for GPCRs. However, the application of directed evolution lead to mutations that improved the stability at the cost of reduced signaling activity.

In the short term, the NMR studies and the complex crystallization trials could be supported by assays that reveal the effects of all individual TM86V mutations on signaling. As previously reported, most mutations appear to have minor, but additive effects on the expression levels and the stability[18]. Rather weak or even neutral mutations in this regard could nevertheless dramatically affect the signaling behavior. For example, the selected mutation R167$^{3.50}$L appears to improve stability only slightly or not at all, but it strongly affects basal and agonist-bound signaling of TM86V-ΔIC3A. Identifying further positions of this type would allow to use more appropriate NTR1 variants in NMR, which simulate the natural activation mechanisms more accurately. Furthermore, a so called “minimal mutant” with similar expression and stability as TM86V but higher signaling levels would also be beneficial for complex crystallization trials, as it may interact more efficiently with G proteins. The generation of such a receptor variant may require the introduction of alternative mutations that further stabilize specific receptor regions, which are not involved in the conformational changes necessary for signaling (e.g. at the TM1-TM2 or at the TM7-TM1 interfaces). These additional mutations may be rationally designed based on the crystal structures and/or using *in silico* free-energy calculations.

It is important to note that stability and signaling are not mutually exclusive. Natural selection itself has generated GPCRs with very high stabilities that seem to act as efficient binary switches, such as rhodopsins, which couple specifically to G_t. Furthermore, *artificial* engineering of certain wild-type GPCRs has also been carried out in the past, so that the receptors interacted preferentially with designed ligands[23; 24]. Similar to my suggestion, some “designer GPCRs” even

activated only G proteins of a specific G_α subtype. In these cases, however, the wild-type receptors already signaled via one specific G_α and only the ligand specificity was changed (usually by abolishing binding to the natural ligand via a point mutation, while maintaining interactions with a specific pharmacological compound). Since the engineered receptors with alternative G protein specificity were based on different wild-type receptors, the amino acid sequences are, however, so different that the G protein specificity determinants remained obscure.

In the long-term, an idea would be to engineer a GPCR, which already couples to various signaling partners as a wild-type, so that each derived variant activates only one specific natural downstream interaction partner. If this was achieved using stable NTR1 variants, the different pharmacological behaviors could be rationalized based on data from the various technologies that are uniquely suitable for this particular model GPCR. NMR, crystallography and DEER spectroscopy (DEER benefits from perdeutaration[23] as well), applied to such receptors, have the potential to answer key questions regarding GPCR activation mechanisms, biased-signaling and G protein specificity. Furthermore, the receptors may be applied in cellular model systems, where they could mimic the effects of perfectly biased GPCR ligands. An improved understanding of GPCR signaling mechanisms and of cellular responses of biased receptors would facilitate the design of ligands that stimulate specific downstream cascades, thus limiting "on-target side effects".

Notably, Matthias Hillenbrand has established a reproducible signaling assay procedure in a 96-well format, where individual G proteins and GPCRs with known absolute protein concentrations can be freely combined. His work also provided data on the stability (or instability) of several dozen heterotrimeric G protein subunit combinations. The availability of all these cloned constructs as viral particles for Sf9 insect cell infection is an excellent basis for the engineering of specific signaling characteristics. While signaling screening procedures are established in this laboratory, evolution methods with directed selection pressures towards signaling are currently not available. Methods of this type may support the above-mentioned project idea and they would likely reveal further details about GPCR activation principles.

4.4. Materials and Methods

4.4.1. Expression in minimal medium (adapted Wagner-lab protocol)

Expression for $^{1}H–^{15}N$ TROSY experiments was carried out in the *E. coli* strain BL21 Tuner using the same pBR322-derived vector as for crystallographic purposes (chapter 2, 3 and 5 [24]). The cells were freshly transformed for each expression and directly cultivated in 5 ml LB

containing 1 % (w/v) glucose and 100 µg/ml ampicillin at 37°C overnight. The saturated culture centrifuged at 4'000 g for 5 min and the cells were resuspended in 500 µl LB. 100 ul of the resuspension was then mixed to 5 ml LB-D_2O containing 1 % (w/v) glucose and 100 µg/ml ampicillin. The culture was incubated for 7 h at 37°C and subsequently centrifuged at 4'000 g for 5 min. All cells were transferred to 50 ml of M9-D_2O minimal medium containing 1 % (w/v) glucose, 100 µg/ml ampicillin and 1 g/l $^{15}NH_4Cl$. The pre-culture was incubated overnight at 37°C. One liter of pre-warmed M9-D_2O minimal medium containing 1 % (w/v) glucose, 100 µg/ml ampicillin and 1 g/l $^{15}NH_4Cl$ was subsequently inoculated to an OD of 0.05. The culture was incubated at 37 °C and induced at an approximate OD of 0.7 by 1 mM IPTG. One hour before induction, the temperature of the culture was lowered to 30°C. The induction period was 16 – 20 h. Cell pellets were stored at -80°C.

4.4.2. NMR experiments on TM86V-ΔIC3 and HTGH4-ΔIC3

Purification of TM86V-ΔIC3B and HTGH4-ΔIC3B for NMR experiments was carried out according to the protocol described in Chapter 2 by means of the non-mutated neurotensin ligand column (pD-NT). The only exception to the aforementioned procedure was that detergent exchange to either 0.05 % (w/v) DDM, 0.2 % (w/v) OGNG or 0.1 % (w/v) DH_7PC was performed during cation exchange chromatography and not at the ligand column step. After the final SEC step (Superdex 200 10/300 GL column) the receptors were concentrated to 100 – 200 µM for $^{1}H–^{15}N$ TROSY experiments, which were conducted at Bruker AvanceIII 600 or 800 MHz magnets at 303, 313 or 323°K, respectively.

4.4.3. Circular dichroism (CD) spectroscopy

Purification of HTGH4-ΔIC3 for NMR experiments was carried out according to the protocol described in Chapter 2 by means of the non-mutated neurotensin ligand column (pD-NT). CD spectroscopy was performed at a Jasco CD Spectropolarimeter J-810 at a protein concentration of 5 µM. Thermal transitions were recorded using purified receptors at a concentration of 5 µM in the buffer of the final SEC run. Cuvettes of 0.1 cm pathlength were used. The secondary structure changes were monitored at a wavelength of 220 nm using a heating rate of 60 K/h. Fitting of the thermal unfolding profiles was carried out assuming a two-state folding mechanism[25].

4.4.4. Nanodisc reconstitution

The best reconstitution results were obtained in the following way at 4°C: After the last purification step (Superdex 200 10/300 GL column), TM86V-ΔICB was frozen in liquid nitrogen for storage at a concentration of 0.53 mg/ml in 10 mM HEPES pH 8, 2 mM DTT, 150 mM NaCl, 0.02 % (w/v) DDM, 30 % (v/v) glycerol and 100 nM NTI (chemically synthesized ligand,

corresponding to the proteolytic cleavage product after ligand column elution). After thawing on ice, the protein was concentrated to less than a tenth of its volume (50 kDa cutoff Amicon Ultra-15), followed by dilution to the original volume using a buffer containing 20 mM TrisHCl pH 7.4, 100 mM NaCl, 0.5 mM EDTA, 0.02 % (w/v) DDM, 250 nM NTI and 2 mM DTT (buffer exchange). Subsequently, the receptor concentration was increased to 100 µM using the same concentrator. A stock concentration of 50 mM of 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) was prepared in a buffer containing 100 mM cholate, 20 mM TrisHCl pH 7.4, 100 mM NaCl and 0.5 mM EDTA. The stock of MSP1D1[21] was prepared at 600 µM in 20 mM TrisHCl pH 7.4, 100 mM NaCl and 0.5 mM EDTA (note that the membrane scaffold protein tends to be insoluble at concentrations above 1 mM).

Using these stock concentrations a 150 µl reconstitution reaction mixture was prepared in an Eppendorf tube by adding 55 µl MSP1D1, 21 µl POPC, 44 µl buffer (20 mM TrisHCl pH 7.5, 100 mM NaCl, 0.5 mM EDTA, 2 mM DTT) and 30 µl TM86V-ΔICB. Note that the receptor was added last after gently mixing the previous components. The final concentrations were 220 µM MSP1D1, 7 mM POPC, 20 µM TM86V-ΔIC3B, 14 mM cholate (CMC: 6 – 14 mM), 0.4 – 4 mM DDM, 20 mM TrisHCl pH 7.4, 100 mM NaCl, 0.5 mM EDTA, 2 mM DTT, 50 nM NTI and trace concentrations of buffer components from the receptor preparation that were not entirely exchanged. The mixture was incubated for 1 – 2 h on ice, followed by dialysis against 3 × 300 ml of a buffer containing 20 mM TrisHCl pH 7.5, 100 mM NaCl, 0.5 mM EDTA, 2 mM DTT and 10 nM NTI (Slide-A-Lyzer Mini Dialysis Unit; 10 kDa cutoff) over a total of 3 days.

The nanodisc reconstitution mixture was diluted 5-fold by SP buffer A containing 20 mM HEPES pH 6.5 and 2 mM DTT. Subsequently it was subjected to a pre-equilibrated SP Sepharose column (0.25 ml bed volume, bench top) using gravity flow. The column was washed by 2 ml SP buffer A and subsequently, the nanodiscs containing receptors were eluted by 4 × 100 µl SP buffer B containing 20 mM HEPES pH 6.5, 2 mM DTT, 250 nM NT1 and 450 mM NaCl.

4.5. References

1. Peddibhotla, S., Hedrick, M. P., Hershberger, P., Maloney, P. R., Li, Y. J., Milewski, M., Gosalia, P., Gray, W., Mehta, A., Sugarman, E., Hood, B., Suyama, E., Nguyen, K., Heynen-Genel, S., Vasile, S., Salaniwal, S., Stonich, D., Su, Y., Mangravita-Novo, A., Vicchiarelli, M., Roth, G. P., Smith, L. H., Chung, T. D. Y., Hanson, G. R., Thomas, J. B., Caron, M. G., Barak, L. S. & Pinkerton, A. B. (2013). Discovery of ML314, a Brain Penetrant Nonpeptidic beta-Arrestin Biased Agonist of the Neurotensin NTR1 Receptor. *ACS Med. Chem. Lett.* **4**, 846-851.
2. Fan, Y., Lai, M. H., Sullivan, K., Popiolek, M., Andree, T. H., Dollings, P. & Pausch, M. H. (2008). The identification of neurotensin NTS1 receptor partial agonists through a ligand-based virtual screening approach. *Bioorg. Med. Chem. Lett.* **18**, 5789-5791.
3. Schlinkmann, K. M., Honegger, A., Tureci, E., Robison, K. E., Lipovsek, D. & Pluckthun, A. (2012). Critical features for biosynthesis, stability, and functionality of a G protein-coupled receptor uncovered by all-versus-all mutations. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 9810-9815.
4. Deupi, X. & Kobilka, B. K. (2010). Energy Landscapes as a Tool to Integrate GPCR Structure, Dynamics, and Function. *Physiology* **25**, 293-303.
5. Manglik A, K. B. (2014). The role of protein dynamics in GPCR function: insights from the beta 2 adrenergic receptor and rhodopsin. *Curr. Opin. Cell Biol.* **27**, 136 - 143.
6. Hofmann, K. P., Scheerer, P., Hildebrand, P. W., Choe, H. W., Park, J. H., Heck, M. & Ernst, O. P. (2009). A G protein-coupled receptor at work: the rhodopsin model. *Trends Biochem. Sci.* **34**, 540-552.
7. Kofuku, Y., Ueda, T., Okude, J., Shiraishi, Y., Kondo, K., Maeda, M., Tsujishita, H. & Shimada, I. (2012). Efficacy of the beta(2)-adrenergic receptor is determined by conformational equilibrium in the transmembrane region. *Nature Communications* **3**.
8. Bokoch, M. P., Zou, Y. Z., Rasmussen, S. G. F., Liu, C. W., Nygaard, R., Rosenbaum, D. M., Fung, J. J., Choi, H. J., Thian, F. S., Kobilka, T. S., Puglisi, J. D., Weis, W. I., Pardo, L., Prosser, R. S., Mueller, L. & Kobilka, B. K. (2010). Ligand-specific regulation of the extracellular surface of a G-protein-coupled receptor. *Nature* **463**, 108-U121.
9. Liu, J. J., Horst, R., Katritch, V., Stevens, R. C. & Wuthrich, K. (2012). Biased Signaling Pathways in beta(2)-Adrenergic Receptor Characterized by F-19-NMR. *Science* **335**, 1106-1110.
10. Kim, T. H., Chung, K. Y., Manglik, A., Hansen, A. L., Dror, R. O., Mildorf, T. J., Shaw, D. E., Kobilka, B. K. & Prosser, R. S. (2013). The Role of Ligands on the Equilibria Between Functional States of a G Protein-Coupled Receptor. *J. Am. Chem. Soc.* **135**, 9465-9474.
11. Kahsai, A. W., Xiao, K. H., Rajagopal, S., Ahn, S., Shukla, A. K., Sun, J. P., Oas, T. G. & Lefkowitz, R. J. (2011). Multiple ligand-specific conformations of the beta(2)-adrenergic receptor. *Nat. Chem. Biol.* **7**, 692-700.
12. Fernandez, C. & Wider, G. (2003). TROSY in NMR studies of the structure and function of large biological macromolecules. *Curr. Opin. Struct. Biol.* **13**, 570-580.
13. Venters, R. A., Farmer, B. T., Fierke, C. A. & Spicer, L. D. (1996). Characterizing the use of perdeuteration in NMR studies of large proteins C-13, N-15 and H-1 assignments of human carbonic anhydrase II. *J. Mol. Biol.* **264**, 1101-1116.
14. Grzesiek, S., Anglister, J., Ren, H. & Bax, A. (1993). C-13 Line Narrowing by H-2 Decoupling in H-2/C-13/N-15-Enriched Proteins - Application to Triple-Resonance 4d J-Connectivity of Sequential Amides. *J. Am. Chem. Soc.* **115**, 4369-4370.
15. Yamazaki, T., Lee, W., Arrowsmith, C. H., Muhandiram, D. R. & Kay, L. E. (1994). A Suite of Triple-Resonance Nmr Experiments for the Backbone Assignment of N-15, C-13, H-2 Labeled Proteins with High-Sensitivity. *J. Am. Chem. Soc.* **116**, 11655-11666.
16. Yamazaki, T., Lee, W., Revington, M., Mattiello, D. L., Dahlquist, F. W., Arrowsmith, C. H. & Kay, L. E. (1994). An Hnca Pulse Scheme for the Backbone Assignment of N-15,C-13,H-2-Labeled Proteins - Application to a 37-Kda Trp Repressor DNA Complex. *J. Am. Chem. Soc.* **116**, 6464-6465.

17. Pervushin, K., Riek, R., Wider, G. & Wuthrich, K. (1997). Attenuated T-2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 12366-12371.
18. Sarkar, C. A., Dodevski, I., Kenig, M., Dudli, S., Mohr, A., Hermans, E. & Pluckthun, A. (2008). Directed evolution of a G protein-coupled receptor for expression, stability, and binding selectivity. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 14808-14813.
19. Bayburt, T. H. & Sligar, S. G. (2010). Membrane protein assembly into Nanodiscs. *FEBS Lett.* **584**, 1721-1727.
20. Borch, J. & Hamann, T. (2009). The nanodisc: a novel tool for membrane protein studies. *Biol. Chem.* **390**, 805-814.
21. Hagn, F., Etzkorn, M., Raschle, T. & Wagner, G. (2013). Optimized Phospholipid Bilayer Nanodiscs Facilitate High-Resolution Structure Determination of Membrane Proteins. *J. Am. Chem. Soc.* **135**, 1919-1925.
22. Inagaki, S., Ghirlando, R. & Grisshammer, R. (2013). Biophysical characterization of membrane proteins in nanodiscs. *Methods* **59**, 287-300.
23. Richard, W., Bowman, A., Sozudogru, E., El-Mkami, H., Owen-Hughes, T. & Norman, D. G. (2010). EPR distance measurements in deuterated proteins. *J. Magn. Reson.* **207**, 164-167.
24. Egloff, P., Hillenbrand, M., Klenk, C., Batyuk, A., Heine, P., Balada, S., Schlinkmann, K. M., Scott, D. J., Schutz, M. & Pluckthun, A. (2014). Structure of signaling-competent neurotensin receptor 1 obtained by directed evolution in Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.* **111**, E655-E662.
25. Privalov, P. L. (1979). Stability of proteins: small globular proteins. *Adv. Protein Chem.* **33**, 167-241.

Chapter 5

Appendices

Content

Chapter 5 93
5.1. Manuscript 95
5.1.1. Abstract 96
5.1.2. Introduction 98
5.1.3. Materials and Methods 99
5.1.3.1 Stabilisation of NTS1 using CHESS 99
5.1.3.2 Stability analysis of 96 individual NTS1 variants from the selected population 100
5.1.3.3 Construct optimization 101
5.1.3.4 Stability comparison of engineered NTS1 variants 101
5.1.3.5 KingFisher saturation binding assays 102
5.1.3.6 KingFisher competition binding assays 103
5.1.4. Results 103
5.1.4.1 CHESS based evolution of apo-state stable NTS1 variants 103
5.1.4.2 Isolating NTS1-H4 from the CHESS selected population 103
5.1.4.3 Comparison of the stability of NTS1-H4 to other NTS1 variants 104
5.1.4.4 Saturation binding of NT to solubilised NTS1-H4 104
5.1.4.5 Competition binding assays using solubilised NTS1-H4 105
5.1.4.6 Sequence of NTS1-H4 105
5.1.5. Discussion 105
5.1.6. References 115
5.2. Abbreviations 118
5.3. Plasmid for large-scale expression of NTR1 variants 120
5.4. Potential applicability of purification principles for other GPCRs 121
5.5. References 122
5.6. Curriculum vitae 123

5.7. Acknowledgments 125

5.1. Manuscript

Title: Improving the apo-state detergent-stability of NTS1 with CHESS for pharmacological and structural studies

Authors: Daniel J. Scott[1,2], Lutz Kummer[1], Pascal Egloff[1], Ross A.D. Bathgate[2] and Andreas Plückthun[1]*

Affiliation: [1]Department of Biochemistry, University of Zurich, 8057 Zurich, Switzerland [2]The Florey Institute of Neuroscience and Mental Health and The Department of Biochemistry and Molecular Biology, The University of Melbourne, Parkville, Victoria 3010, Australia.

* Corresponding author.

Address: Department of Biochemistry,
University of Zurich,
Winterthurerstrasse 190,
8057 Zurich, Switzerland.
Phone: +41 44 635 55 70. Fax: +41 44 635 57 12
E-mail: plueckthun@bioc.uzh.ch

5.1.1. Abstract

The largest single class of drug targets is the G Protein-Coupled Receptor (GPCR) family. Modern high-throughput methods for drug discovery require working with pure protein, but this has been a challenge for GPCRs, and thus the success of screening campaigns targeting soluble, catalytic protein domains has not yet been realized for GPCRs. Therefore most GPCR drug screening has been cell-based, whereas the strategy of choice for drug discovery against soluble proteins is HTS using purified proteins coupled to structure-based drug design. While recent developments are increasing the chances of obtaining GPCR crystal structures, the feasibility of screening directly against purified GPCRs in the unbound state (apo-state) remains low. GPCRs exhibit low stability in detergent micelles, especially in the apo-state, over the time periods required for performing large screens. Recent methods for generating detergent-stable GPCRs, however, offer the potential for researchers to manipulate GPCRs almost like soluble enzymes, opening up new avenues for drug discovery. Here we apply Cellular High-throughput Encapsulation, Solubilization and Screening (CHESS) to the neurotensin receptor 1 (NTS1) to generate a variant that is stable in the apo-state when solubilized in detergents. This high stability facilitated the crystal structure determination of this receptor and also allowed us to probe the pharmacology of detergent-solubilized,apo-state NTS1 using robotic ligand binding assays. NTS1 is a target for the development of novel antipsychotics and thus CHESS-stabilized receptors represent exciting tools for drug discovery.

Keywords: G Protein-Coupled Receptor, directed evolution, CHESS, stabilization, encapsulation, detergent, apo-state, thermostability.

Abbreviations: GPCR, G protein-coupled receptor; IMP, integral membrane protein; CHESS, cellular high-throughput encapsulation solubilization and screening; NT, neurotensin peptide; FACS, fluorescence-activated cell sorting; EDTA, ethylenediamine tetraacetic acid; DDM, n-dodecyl-β-D-maltopyranoside; DM, n-decyl-β-D-maltopyranoside; OG, n-octyl-β-D-glucopyranoside; CHAPS, 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate; CHS, cholesteryl hemisuccinate; HTG, n-heptyl-β-D-thioglucopyranoside; HTS, high-throughput screening; PBS, phosphate buffered saline; sfGFP, super-folder green fluorescent protein.

Highlights:

- NTS1 was stabilized in harsh, short-chain detergents using CHESS, generating NTS1-H4
- NTS1-H4 exhibited a 26.8°C improvement in apo-state thermostability
- Apo-state NTS1-H4 could be stored for hours without significant loss of activity
- Solubilised apo-state NTS1-H4 was used for in vitro, high-throughput binding assays
- The crystal structure of NTS1-H4 in short-chain detergent was recently solved

5.1.2. Introduction

GPCRs are located in the cell membranes of all human cell types where they serve to detect and transduce extracellular signals into intracellular signaling pathways. The GPCR gene family is the largest in the human genome and encodes approximately 850 different receptors that sense and respond to a huge variety of stimuli including neurotransmitters, metabolites, hormones and environmental stimuli such as light, tastes and smells [1]. This diverse array of stimuli is a testament to the evolutionary success of the protein architecture of GPCRs, made up of seven transmembrane spanning domains, which is maintained throughout the family despite low sequence homology. Upon activation, GPCRs couple with, and stimulate intracellular G-proteins to initiate cellular signaling pathways.

Because of the location of GPCRs on the surface of cells and their involvement in many, if not most, physiological pathways, GPCRs are the major class of drug targets in the human body [2]. Conversely, less than 10% of the GPCR family is currently targeted by prescription drugs [2]. This discrepancy is primarily due to the lack of knowledge about how molecules interact with and activate GPCRs at the molecular level, such that a true molecular design of specific agonists and antagonists has not been possible. Additionally, for a great number of receptors, neither the natural ligand nor the function have been elucidated ("orphan receptors"). Most drugs have thus come from cellular screening of the known receptors.

Modern drug discovery techniques for targeting soluble enzymes for example, have higher success rates based on improved in vitro screening assays and the parallel application of surface plasmon resonance (SPR) or nuclear magnetic resonance spectroscopy (NMR) based fragment screening in conjunction with structure based lead optimization [3]. The hurdle for structural, mechanistic and in vitro drug screening studies of GPCRs is that to apply a similar workflow to GPCRs, the receptors must be solubilized in detergents and purified. However, GPCRs typically exhibit low stability in detergent micelles, especially when a sample is required to be stable in the apo-state for many hours to facilitate in vitro binding assays and fragment screening using biophysical methods.

Recent progress in obtaining crystal structures of GPCRs [4-19] will undoubtedly aid in the computational optimization of drug leads. Most of these structures were solved as fusion proteins, with T4-lysozyme replacing intracellular loop 3 [5, 20], which acts as a rigid scaffold that promotes the formation of crystal contacts in lipidic-cubic phase crystallization trials. The fusion strategy is necessary so that sufficient protein surface area is displayed outside of the lipid bilayer, because for most unmodified GPCRs, virtually all of the protein is embedded within the bilayer and thus unable to contribute to crystal contact formation. This technique does not significantly increase the stability of the receptor in the solubilised state [20, 21] and, because

the receptors are reconstituted into insoluble media, this strategy is not useful for direct high-throughput screening for interacting molecules. Furthermore, such a fusion disallows interaction between the GPCR and G proteins so that signaling as a readout is not an option.

A pioneering approach to making GPCRs more accessible to structural, biochemical and biophysical methods is to stabilise the receptors by introducing thermostabilizing mutations. Stabilising mutations have been identified using semi-rational or alanine-scanning and screening approaches [5, 22-26], and with directed evolution methods [27-31]. Stabilized GPCRs can be successfully applied to crystallization [7, 14, 19, 32-35], robotic in vitro binding assays in the solubilised form [31] and fragment screening using biophysical methods [36, 37]. Generally, these studies require the receptor to be purified in the presence of a ligand to stabilise the receptors during the time needed for purification and assay setup. This necessitates extensive washing to remove bound ligand before assays can be conducted [36], and the instability of these receptors in the apo-state may limit the time that samples can be probed to unrealistic intervals. To enable reliable measurements of GPCR samples over the time scales required for high-throughput screening (HTS) assays or NMR-based fragment screening, receptors are required that are stable for many hours, preferably in the apo-state. Here we use the Cellular High-throughput Encapsulation, Solubilization and Screening (CHESS) method to evolve neurotensin receptor 1 (NTS1) mutants that meet these requirements. The crystal structure of one of the resultant receptors was recently solved in a detergent-solubilized form [19]. Here we demonstrate that the long-term, apo-state stability of these CHESS-stabilized NTS1 variants makes them suited to HTS-compatible ligand binding assays using isolated receptors in detergent micelles.

5.1.3. Materials and Methods

5.1.3.1 Stabilisation of NTS1 using CHESS

E. coli strain DH5α was transformed with the StEPM303 library and receptor expression was induced as described previously [29, 31]. 1.75 x 1010 cells from the expression culture were encapsulated with one layer of chitosan and 1 layer of alginate as described previously [31]. For the initial selection round, the capsule population was exposed to PBS pH 7.4, 1 mM EDTA and 1.7% DM (termed PBS-E(DM)) for 3 h at 20°C with vigorous shaking without ligand, followed by 2 h at 20°C in the presence of 20 nM BODIPY FL-labeled NT(8-13) (FL-NT). Capsules were washed twice in PBS-E(DM) and subjected to FACS using a FACSAria III cell sorter (BD Biosciences), where capsules were sorted that exhibited fluorescence placing them in the top 0.5 – 1% of the population, resulting in the collection of 50,000 capsules. Genetic information was recovered from sorted capsules by PCR amplification using NTS1 specific primers after incubation in an ultrasonic water bath for 5 minutes. For the second round of selection, the

amplified receptor genes were re-cloned, proteins expressed and the cells encapsulated as above. The capsule population was first treated with PBS-E(DM) for 3 h at 20°C, followed by addition of 20 nM FL-NT for 1 h at 20°C, before the capsules were collected by centrifugation and resuspended in PBS-E containing 2% octyl-β-D-glucopyranoside (OG) (PBS-E(OG)) and 20 nM FL-NT at 4°C. Capsules were washed once in 20 nM FL-NT in PBS-E(OG) to promote efficient detergent exchange before being incubated for 2 h in PBS-E(OG) with ligand at 4°C. Capsules were washed twice in PBS-E(OG) and subjected to FACS as above, resulting in the collection of 38,000 capsules. These clones were isolated and re-cloned as above for a third round of selection. In the third round of selection, the capsules were exposed to PBS-E(DM) for 3 h at 20°C, followed by addition of 20 nM FL-NT for 1 h at 20°C, before the capsules were collected by centrifugation and resuspended in PBS-E containing 2% heptyl-β-D-thioglucopyranoside (HTG) (PBS-E(HTG)) and 20 nM FL-NT at 4°C. Capsules were washed once in 20 nM FL-NT in PBS-E(HTG) to promote efficient detergent exchange before being incubated for 25 h at 4°C. After this step, capsules were washed twice in PBS-E(HTG) and sorted with FACS as above, resulting in the collection of 20,000 capsules. The DNA encoding these clones was isolated and re-cloned into an expression vector containing a C-terminal sfGFP fusion, as in [31]. The CHESS workflow is depicted in Figure 5.1.

5.1.3.2 Stability analysis of 96 individual NTS1 variants from the selected population

Forty-seven individual colonies derived from the capsules sorted in the 3rd round of CHESS were picked and used to inoculate 1.2 ml cultures of LB broth containing 100 μg/ml ampicillin and 1% glucose in a 96-deep-well plate. As a control, a colony of cells transformed with a previously stabilised NTS1 variant (C7E02) [29] was also picked. Cultures were grown for 16 h at 37°C before being used to inoculate 48, 5-ml-expression cultures in 24-deep-well plates. Receptors were expressed as described previously at 20°C for 20 h [31]. The initial cultures were centrifuged and the plasmid DNA isolated. After expression, cell pellets were solubilised in 1 ml PBS-E(DM) for 2 h at 20°C and the cell debris pelleted with centrifugation at 5000 g for 20 min. 0.8 ml of the supernatant was transferred to a 96-well KingFisher plate (Thermo Scientific) containing 20 μg of streptavidin T1 Dynabeads (Life Technologies) per well. Samples were then robotically manipulated using a KingFisher Flex robot (Thermo Scientific) as described previously [31], including binding of 20 nM HL-NT for 1 h, exchanging the detergent to 2% HTG for 1.5 h, washing away unbound ligand for 5 min and eluting the beads into 0.25 ml PBS-E(HTG) for analysis of ligand binding in a fluorescent plate reader (Tecan M1000) and with flow cytometry (Partec CyFlow Space). 100,000 beads from the binding assay were measured with flow cytometry, with the average fluorescence intensity of single-sized beads presented in Figure 5.2B upon 638 nm laser excitation and emission at 675 nm (20 nm bandpass).

5.1.3.3 Construct optimization

Based on the stability screen, the most promising stabilized variant, number 47 (termed NTS1-H4) was sequenced. Before further analysis with fluorescent binding assays, the following changes were made to the receptor. A mutation in the conserved E/DRY motif was reverted (L167R), a potential human rhinovirus 3C protease site was removed from intracellular loop 3 (Q274A), alanine 342 in extracellular loop 3 was reverted to the naturally occurring phenylalanine (A342F) and all exposed cysteines were mutated to either alanine or serine (C278A, C386A, C388A, C417S). The resultant receptor was termed NTS1-H4(BM1) and the encoding gene was synthesized by Genscript.

5.1.3.4 Stability comparison of engineered NTS1 variants

Stability measurements of selected NTS1 receptor variants were performed as described previously [28, 31]. NTS1 receptor variants were expressed with a C-terminal sfGFP-Avi-tag fusion. A cell pellet corresponding to a 2.5 ml expression culture was used for one single measurement reporting functionally folded receptor as determined by a ligand binding assay. Receptors were solubilized in solubilization buffer (PBS pH 7.4, 1% (w/v) DDM, 0.5% (w/v) CHAPS, 0.1% (w/v) CHS, 1 mM EDTA, Complete protease inhibitors (Roche), 40 µg/ml deoxyribonuclease I (Roche), 1 mg/ml lysozyme). Solubilization was performed at 4°C for 2 h. Cell debris was removed by centrifugation and the supernatant was exposed to streptavidin-coated paramagnetic beads. Solubilized receptor variants were allowed to bind to the beads for 1 h at 4°C before being transferred in 96-well plates for subsequent manipulation with a KingFisher Flex magnetic particle processor. Receptor-coated beads were subjected to detergent solution, PBS-E(DCC) (PBS pH 7.4, 1 mM EDTA, 1% (w/v) DDM, 0.5% (w/v) CHAPS, 0.1% (w/v) CHS) with or without 20 nM Hilyte-647 labeled neurotensin 8-13 (HL647-NT). Non-specific binding was determined by adding excess (1 µM) unlabeled NT8-13 competitor to the binding solutions. After a further 0 h or 18 h exposure to DDM, beads were washed twice with PBS-E(DCC) and those exposed to detergent in the unbound state were transferred to solutions containing 20 nM HL647-NT with or without competitor for 1.5 h. Receptor-coated beads were washed twice in PBS-E(DCC) and transferred to black 96-well microplates (Greiner) in a final volume of 100 µl. HL647-NT and sfGFP fluorescence levels were measured in each well using an M1000 dual monochromator fluorescence plate reader (Tecan) with excitation at 630 nm for HL647-NT and 488 nm for sfGFP. Fluorescence emission for HL647-NT was measured at 680 nm and for sfGFP at 530 nm.

Thermal stability measurements in PBS-E(DCC) of NTS1 receptor variants were essentially performed as described for the analysis of detergent stability (see above). After exposure of solubilized receptor variants to streptavidin-coated paramagnetic beads, receptor-coated beads

were resuspended in PBS-E(DCC) with or without ligand (or competitor). Beads were transferred to 96-well PCR plates and exposed to different temperatures for 30 min in a gradient PCR cycler (Biometra). After heat treatment, receptor-coated beads that were heated in the absence of ligand were incubated with PBS-E(DCC) containing 20 nM HL647-NT or competitor for 1.5 h. Beads were washed twice in PBS-E(DCC) and transferred to black 96-well microplates and fluorescence intensities of HL647-NT and sfGFP in each well were determined as above. The data were analyzed by nonlinear regression fitting with GraphPad Prism.

5.1.3.5 KingFisher saturation binding assays

NTS1-H4(BM1) was expressed in 400 ml cultures for 20–24 h at 20 °C, the cells harvested with centrifugation and the pellet resuspended in 10 ml 50mM HEPES (pH 7.8) containing 200 mM NaCl, 10 mM EDTA, 1 mg/ml chicken lysozyme, 10 U/ml DNAse, 1.7% DM and 0.5% CHAPS. Cells were solubilized at 20°C with vigorous shaking for 3 h. Cell debris was removed with centrifugation at 15,000 g for 10 minutes at 4°C. The supernatant containing solubilized receptor was then incubated with streptavidin-coated paramagnetic beads, 2.5 µg of beads per ml of culture, at 4°C for 1 h to immobilize the biotinylated receptor onto the beads. Beads were collected with centrifugation at 5,000 g for 5 min and resuspended in assay buffer (50 mM HEPES (pH 7.8), 200 mM NaCl, 10 mM EDTA and 0.1% DM) at a final concentration of 100 µg beads per ml. 20 µg of beads were then added to 48 wells of a deep well KingFisher plate (plate 1) (Thermo Scientific). A concentration series of Alexa-647 labelled neurotensin 8-13 (A647-NT) was made in assay buffer and 1 ml of each added to several wells of a separate KingFisher plate (plate 2). For each A647-NT concentration, a separate solution was made containing an excess of unlabeled neurotensin 8-13 (10 µM), with 1 ml of each being aliquoted into designated wells of plate #2. 200 µl and 100 µl of assay buffer was added to 48 wells of another two KingFisher plates, plate #3 and plate #4 respectively. A KingFisher 96 magnetic particle processor was used to automatically perform the following steps at 4°C: the beads were captured from plate #1 and transferred to plate #2, beads and ligand solutions were mixed for 2 h at 4°C, beads were transferred to plate #3 and washed for 1.5 min before being transferred to plate #4. Beads were transferred with a multichannel pipette from KingFisher plate 4 to a Greiner non-binding black 96-well plate. The sfGFP (excitation filter 485 nm, band pass 12 nm, emission filter 520 nm, band pass 10 nm) and Alexa-647 (excitation filter 640 nm, band pass 10 nm, emission filter 670 nm, band pass 10 nm) signals from each well were measured using an Omega Polarstar plate reader (BMG Labtech). To determine if ligand depletion was occurring during the binding incubation, 50 µl of the A647-NT solutions in KingFisher plate #2 were transferred to a Greiner non-binding fluorescence 96-well plate and the Alexa-647 signals of the wells compared. Data were analysed with Graphpad Prism, with curves fitted using the one site – total and nonspecific binding equations.

5.1.3.6 KingFisher competition binding assays

Receptor-coated bead samples were prepared and aliquoted into KingFisher plate #1 as described in 2.5. In plate #2, concentration series of the various competitors were added, all in assay buffer supplemented with 2 nM A647-NT, 0.5 ml per well. Competitors included SR 48692 (Tocris Biosciences and Sigma Aldich), SR 142948 (Tocris Biosciences), neurotensin 8-13 (Sigma Aldrich) and neurotensin 1-12 (synthesized by GL Biochem, Shanghai, China). The binding assays were performed using a KingFisher 96 robot and Omega Polarstar plate reader as in 2.5. Data were analysed using Graphpad Prism with curves fitted using the one site – fit Ki equation. Data from 3 separate experiments were pooled and the estimated Kd values for A647-NT calculated in 2.5 were used to fit the competition curves.

5.1.4. Results

5.1.4.1 CHESS based evolution of apo-state stable NTS1 variants

To evolve an NTS1 variant that could be purified in short-chain detergents for X-ray crystallization and exhibit apo-state and long-term stability in the solubilized form, we used the highly diverse StEPM303 library of NTS1 mutants and selected using CHESS following the strategy outlined in Figure 5.1. In each selection round, the encapsulated GPCRs were solubilized for three hours, at 20°C, in the absence of ligand to place selective pressure on apo-state stability. In the second and third rounds of selection, the encapsulated receptor population was exposed to the short-chain detergents OG and HTG respectively, placing selective pressure on the population for stability in detergents that are suitable for vapor diffusion crystallization of GPCRs. Finally, long-term stability was selected for by leaving the third generation of the evolving population of GPCRs in HTG for 25 h before selecting out the most stable clones with FACS.

5.1.4.2 Isolating NTS1-H4 from the CHESS selected population

The genes encoding the CHESS-selected NTS1 variants were isolated from the sorted capsules and cloned into an expression vector comprising C-terminal sfGFP and avi-tag (for in vivo biotinylation) fusions. The stability of 47 receptor variants was assayed by testing their ability to bind a fluorescently labeled neurotensin ligand after solubilization in HTG for 2 h and 98 h (Fig. 5.2). In this assay, solubilized receptors were captured on streptavidin-coated beads and exposed to fluorescently labelled neurotensin 8-13 (HL-NT). The ability of each clone to bind ligand at the given time points was determined by washing the beads and measuring the amount of HL-NT bound to the beads with a fluorescent plate reader (Fig. 5.2A) and a flow cytometer (Fig. 5.2B). Of these receptors, clone 47 (NTS1-H4) was selected for further analysis

because of its ability to bind high levels of NT after 2 h and 98 h solubilized in HTG, when measured with both methods.

5.1.4.3 Comparison of the stability of NTS1-H4 to other NTS1 variants

To determine whether NTS1-H4 exhibited favorable long-term and apo-state stability, we compared the stability of this receptor to wild-type rat NTS1, a thermostabilized variant produced through systematic mutation by Shibata et al. [26] (NTS1-7m), and a highly optimized variant generated through bacterial display and systematic mutation by Schlinkmann et al. (TM86V) [29, 31]. Each receptor was expressed in E. coli with an sfGFP-avi-tag C-terminal fusion, resulting in the production of fluorescent, biotinylated receptor. Unmodified NTS1 is not stable in detergents such as DM, OG and HTG, so to enable comparison across the receptor variants, we solubilized the cells in a mild detergent mix (PBS-E(DCC)). Solubilized receptors were immobilized on streptavidin-coated beads and subjected to a range of temperatures for 30 minutes in the presence (Fig. 5.3A) or absence (Fig. 5.3B) of fluorescently labelled neurotensin to generate thermostability curves. Such curves are commonly used to rank the stabilities of engineered GPCRs, and the temperature at which only half the receptor proteins are able to bind ligand is referred to as the apparent melting temperature (Tm). NTS1-H4 exhibited the highest thermostability of all the variants when heated in both the presence (Tm = 57°C) and absence (Tm = 48.1°C) of excess neurotensin (Fig. 5.3A-B). Thus, the thermostability of NTS1-H4 in this harsh detergent is improved by 21.6°C in the bound state and 26.8°C in the apo-state compared to unmodified NTS1.

The relative amount of folded receptor in each of the samples was monitored after 3 h in the apo-state and either 21 h in the apo-state or with HL-NT bound at 4 °C. The relative levels of folded receptor were calculated by measuring the ratio of bound HL-NT fluorescence (level of folded protein) to the sfGFP fluorescence (total protein) in each sample at various time points. As expected, unmodified NTS1 displayed the lowest level of folded receptor under each condition, closely followed by NTS1-7m (Fig. 5.3C). Interestingly, for all the receptor samples, there was no significant difference in the relative NT binding levels after a 21 h incubation at 4°C in the presence of HL-NT. There was a striking difference however, when the solubilized receptors were incubated for 21 h at 4 °C in the absence of NT (Fig. 5.3C). Under these conditions, no binding of NT could be detected on NTS1, while NT binding to NTS1-7m and TM86V was decreased by 72% and 63% respectively, whereas only a 22% decrease in ligand binding was seen for NTS1-H4.

5.1.4.4 Saturation binding of NT to solubilised NTS1-H4

The high stability exhibited by NTS1-H4 in the apo-state indicated that this variant could be used to probe the binding of ligands to solubilized, isolated receptor preparations in a low-

cost high-throughput compatible way. To demonstrate this, we expressed and captured NTS1-H4-sfGFP onto magnetic beads and using a KingFisher magnetic particle processor and fluorescent plate reader, conducted saturation binding assays using Alexa647-NT (Fig. 5.4A). The ligand binding step was conducted for 2 h at 4°C in 1 ml of solution, which resulted in no significant ligand depletion, even at only 100 pM A647-NT (Fig. 5.4B). Non-specific binding was determined by measuring the binding of increasing concentrations of A647-NT in the presence of 1 μM unlabeled NT8-13. To control for differences in bead loading, sfGFP fluorescence was measured in each well and the specific binding was calculated as a ratio of A647-NT fluorescence to sfGFP fluorescence. Fitting the resultant data enabled us to estimate the Kd of A647-NT binding to NTS1-H4 at 0.65 nM ± 0.13.

5.1.4.5 Competition binding assays using solubilised NTS1-H4

To probe the binding of unlabeled ligands to solubilized NTS1-H4, competition binding assays were performed in the same robotic manner, using 2 nM A647-NT as the labelled ligand (Fig. 5.4C). Unlabeled agonists neurotensin 8-13 (NT8-13) and neurotensin 1-12 (NT1-12), along with antagonists, SR48692 and SR142948, were able to compete with A647-NT in a dose-dependent manner (Fig. 5.4C). Using the estimated Kd value for A647-NT from the competition binding experiments we were able to fit Ki values for NT8-13, NT1-12, SR48692 and SR142948 respectively (Table 5.1).

5.1.4.6 Sequence of NTS1-H4

NTS1-H4 was sequenced and found to contain 25 amino acid substitutions over wild-type rat NTS1 including; S83G, A86L, T101R, H103D, H105Y, L119F, M121L, E124D, R143K, D150E, A161V, R167L, R213L, V234L, K235R, V240L, I253A, I260A, N262R, K263R, H305R, C332V, T354S, F358V, and S362A. Of these, H103D, H105Y, A161V, R213L, V234L, H305R and S362A are derived from the parental receptor in the StEPM303 library (NTS1-DO3) and thus the other 18 mutations were acquired during the selection outlined in Figure 5.1. Figure 5.5 indicates the positions of these mutations on the crystal structure of NTS1-H4 [19].

5.1.5. Discussion

The critical advantage of CHESS over other techniques is that millions of GPCR-mutant-containing microcapsules can be screened directly for the desired stability properties within an hour using FACS, Here we panned the StEPM303 library [29] for NTS1 mutants that were stable in short-chain detergents that are desirable for crystallization. In addition, we selected for stability over 24 h in such detergents, as well as receptors that exhibiting apo-state stability. We

believe such properties to be highly desirable for the application of the experimental methods commonly used for SBDD.

The encapsulated cells were resistant to HTG over these time periods, and genes encoding highly stable NTS1 mutants could be isolated from the sorted capsules that exhibited high fluorescent ligand binding (Fig. 5.2). The clone that was able to bind the most fluorescent ligand after 2 h and 98 h solubilized in HTG, when measured using 2 different instruments, was termed NTS1-H4. NTS1-H4 exhibited all of the properties we were seeking, including high thermostability when heated in the presence and absence of ligand (Fig. 5.3A-B), and long-term, apo-state stability in detergent (Fig. 5.3C).

Because NTS1 has been stabilized using several methods, we were able to directly compare the resultant receptors with the wild-type protein. NTS1-7m was stabilised using the systematic mutagenesis approach [26] and TM86V was derived through a combination of bacterial display for high expression and systematic mutagenesis [29]. Using a detergent mixture known to maintain unmodified rNTS1 in a stable state under mild conditions [38, 39], thermostability assays were performed on the 4 receptors (Fig. 5.3A-B). When heated in both the presence and absence of ligand, NTS1-H4 was the most thermostable, followed by TM86V, NTS1-7m and wild type. Compared to rNTS1, NTS1-H4 exhibited a 21.6 °C and 26.8 °C improvement in thermostability in detergent HTG when heated in the bound and apo-state respectively, the most stable by over 10 °C. NTS1-H4 also exhibited the highest stability of the 4 receptors when incubated in this harsh detergent for over 21 h in the presence or absence of bound ligand (Fig. 5.3C).

This high level of apo-state stability meant that we could successfully use bacterially expressed NTS1-H4, captured onto magnetic beads, in robotic fluorescence-based saturation and competition binding assays (Fig. 5.4). From fitting the saturation binding curves the Kd of the fluorescently labelled neurotensin analogue, Alexa-647-labelled neurotensin 8-13 (A647-NT), could be estimated to be 0.65 nM ± 0.13, which compares well to Ki values reported for unlabeled versions of this peptide using NTS1 expressing cells and tissues [40]. The high signal-to-noise ratio in the saturation binding assays using this labelled peptide enabled us to conduct competition binding assays using a sub-saturating concentration of A647-NT (2 nM). Ki values for the binding of unlabeled NT8-13, NT1-12, SR48692 and SR142948 to NTS1-H4 were calculated from the resultant curves (Fig. 5.4C and Table 1). Calculated Ki values compared well to the literature values for NT8-13 [40], NT1-12 [41] and SR142948 [42], where cell culture and tissue preparations were used for NTS1 competition binding assays. This demonstrated that our engineered, bacterially expressed and detergent solubilized GPCR exhibited native-like ligand binding behavior.

Interestingly, the antagonist SR48692 was less potent at competing A647-NT binding to solubilized NTS1-H4, compared to wt NTS1 which was tested in cell-based binding studies, by a factor of between 2.7 and 44 fold, depending on the values cited in the literature [42-44]. The affinity of SR48692 to the thermostabilised NTS1 variant NTS1-GW5 was reduced even by 130-200 fold when measured in insect cell membranes, whereas binding of the agonist to this receptor was unaltered [14]. NTS1-H4 and NTS1-GW5 share only one homologous stabilizing mutation at position 358 in TM7, which is a substitution from Phe to Ala in NTS1-GW5 and to Val in NTS1-H4. Mutation of this position to alanine has been reported to reduce the affinity of SR48692 for NTS1 but not that of neurotensin [44] and result in a constitutively active receptor with respect to inositol phosphate production [45], suggesting that the lower affinity of SR48692 for both NTS1-H4 and NTS1-GW5 may be a result of the mutation at position 358. Overall, engineering NTS1-H4 to be detergent-stable in the apo-state has enabled us to pharmacologically characterize the orthosteric binding site of a detergent-solubilized GPCR using low cost, high-throughput compatible assays.

Thermostabilization of NTS1 has so far resulted in the determination of 5 structures of different receptor variants or constructs. NTS1-GW5 was engineered using systematic mutagenesis [46], and the crystal structure of this receptor, produced in insect cells, was determined using the T4-lysozyme fusion and lipidic cubic phase crystallization approach [14]. The crystal structures of two NTS1-TM86V constructs, derived from directed evolution using bacterial display [29], were recently solved in detergent using vapor diffusion crystallization [19]. In the same study, the structures of two more receptors were determined, NTS1-G7 [31] and NTS1-H4 (engineered in this manuscript), which were stabilised using CHESS. Of note is that the structures of NTS1-TM86V, NTS1-G7 and NTS1-H4 were the first GPCR structures solved from bacterially expressed protein. Furthermore, the direct selection of NTS1-G7 and NTS1-H4 for stability in short-chain detergents allowed these receptors to be crystallised in short-chain detergents using vapor diffusion crystallography without any systematic mutational optimization. TM86V was also crystallized in short chain detergents, but this receptor was generated after further systematic mutational optimization of clones selected using bacterial display which were very stable in milder detergents, but less stable in short-chain detergents [29].

A thorough comparison of the 5 structures can be found in Egloff et al. [19]. While NTS1-H4 is the most thermostable variant described to date, it also contains the most mutations. In Figure 5, the CHESS selected mutations in NTS1-H4 are highlighted in the crystal structure. Selected mutations were relatively evenly distributed across transmembrane domains (TMs) 1-3 and TMs 5-7, with the most mutations occurring in TM5 (6 substitutions) (Fig. 5.5). NTS1-7m, NTS1-GW5, TM86V, NTS1-G7 and NTS1-H4 only share homologous stabilizing mutations at

positions 86 (A86L) and 358 (F358A or V), indicating that stabilization can be achieved through a variety of mutational modes.

In addition to structure determination, purified thermostabilized GPCRs have also proven to be useful tools for biophysical analysis of ligand binding and fragment screening [36, 37, 47]. For these types of studies it is highly beneficial to have apo-receptor samples that are stable in detergents over the long time periods needed to measure binding kinetics, and that won't denature upon ligand dissociation. The engineering of NTS1-H4 is a demonstration of the utility of CHESS for evolving receptors with the properties needed to conduct biochemical and biophysical experiments on solubilized, purified GPCRs.

Table 5.1 Binding constants derived from competition binding experiments using solubilized NTS1-H4

Ligand	Fitted Constant	nM	SEM	Figure
NT8-13	K_i	0.34	0.09	5C
NT1-12	K_i	306	77	5C
SR48692	K_i	87	27	5C
SR142948	K_i	0.5	0.1	5C

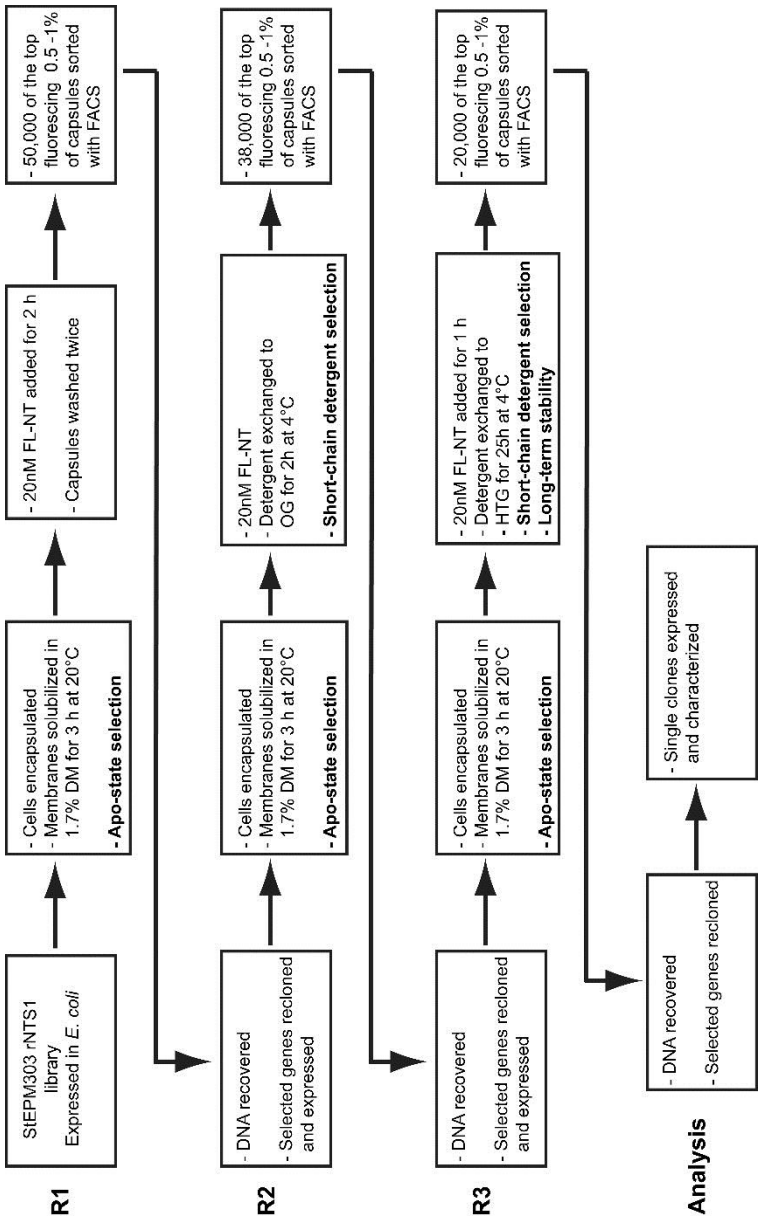


Figure 5.1 CHESS based selection workflow for generating NTS1 mutants that were resistant to short-chain detergents, apo-state solubilization and long term incubations in the solubilized state. CHESS rounds are indicated by R1-3. For a diagrammatic representation of the CHESS method, see Scott et al. [31].

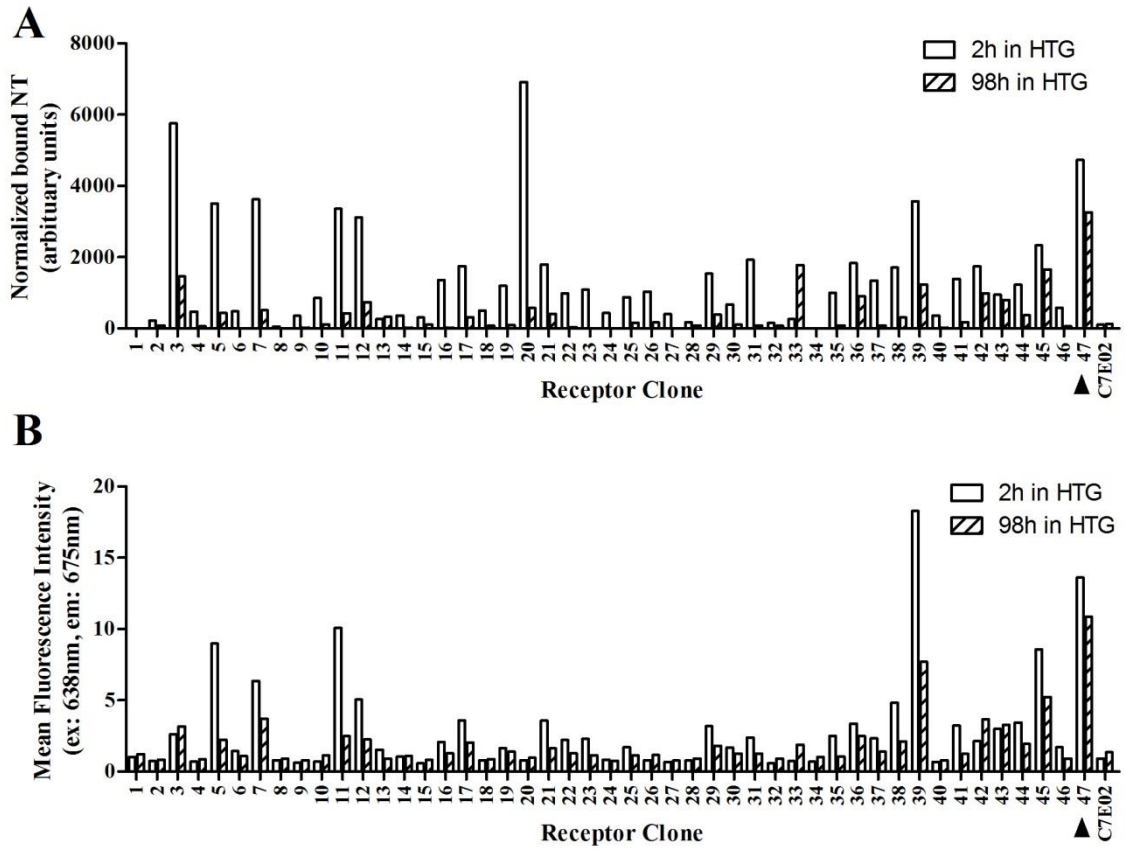


Figure 5.2 Single clone analysis of 47 selected NTS1 mutants after 3 rounds of CHESS. Detergent-solubilized, biotinylated receptors were captured onto streptavidin-coated magnetic beads. The ability of the receptor-coated beads to specifically bind HL-NT after 2 h (open columns) or 98 h (striped columns) in the short-chain detergent HTG was measured using a fluorescent plate reader (A), or an analytical cytometer (B). Clone 47, indicated with arrows, was termed NTS1-H4.

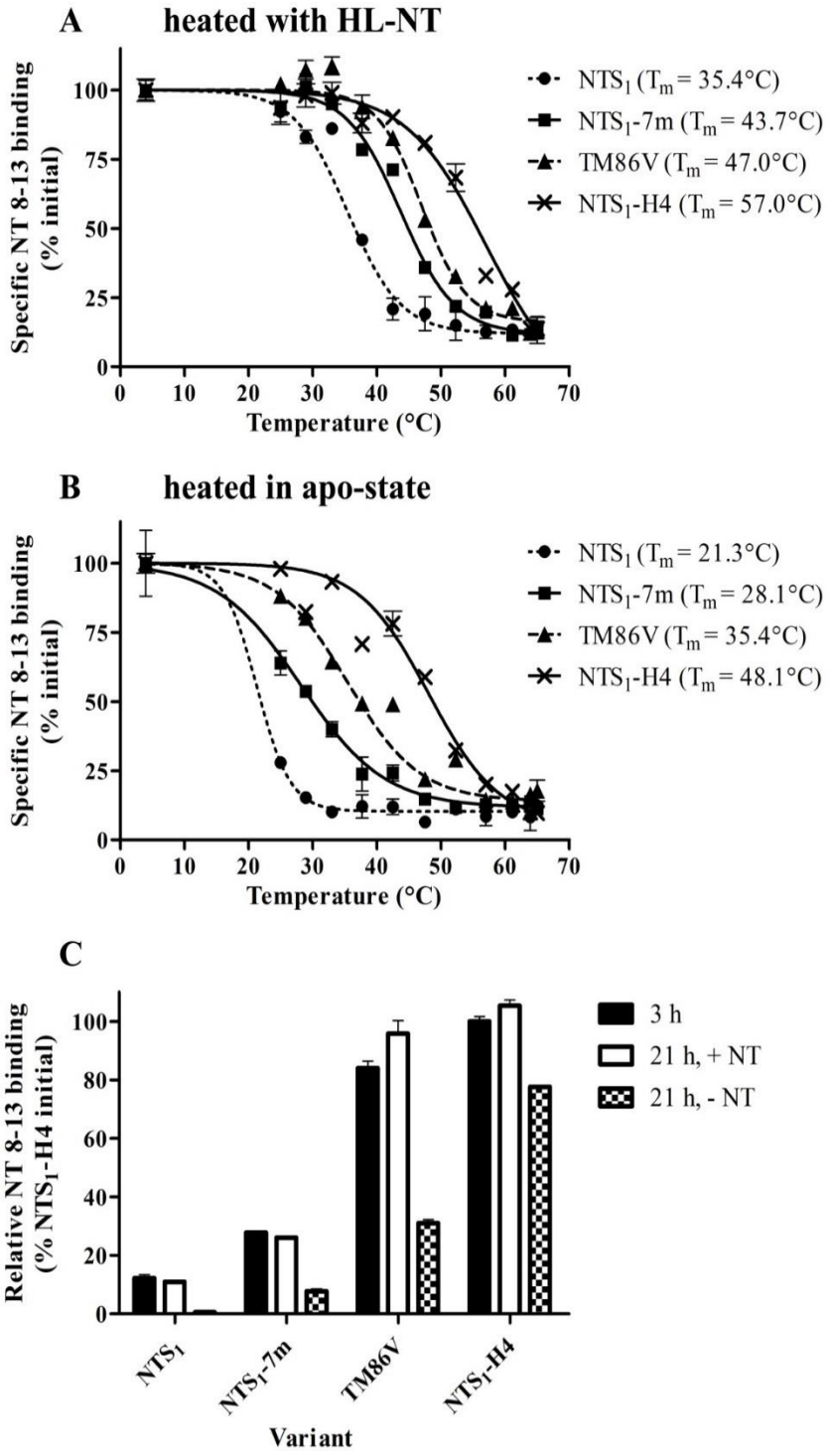


Figure 5.3 The stabilities of thermostabilized NTS1 mutants generated using different methods were compared to wild type NTS1. The thermostabilities of wild-type rat NTS1 (NTS1, filled circles with dotted lines), NTS1-7m (filled squares with solid lines), TM86V (filled triangles with dashed lines) and NTS1-H4 (crosses with solid lines) were measured with the receptors heated in the presence of HL-NT (A) or in the apo-state (B). Apparent melting temperatures (Tm) were determined with non-linear regression and are displayed in parentheses next to the figure keys. (C) Long-term stability of the 4 receptors was assayed by measuring the relative levels of HL-NT binding 3 h after solubilization (filled columns), 21 h after solubilization, 18 h of which the receptors were incubated with HL-NT (open columns), or 21 h after solubilization incubated in the apo-state (chequered columns). The receptors were incubated at 4 °C.

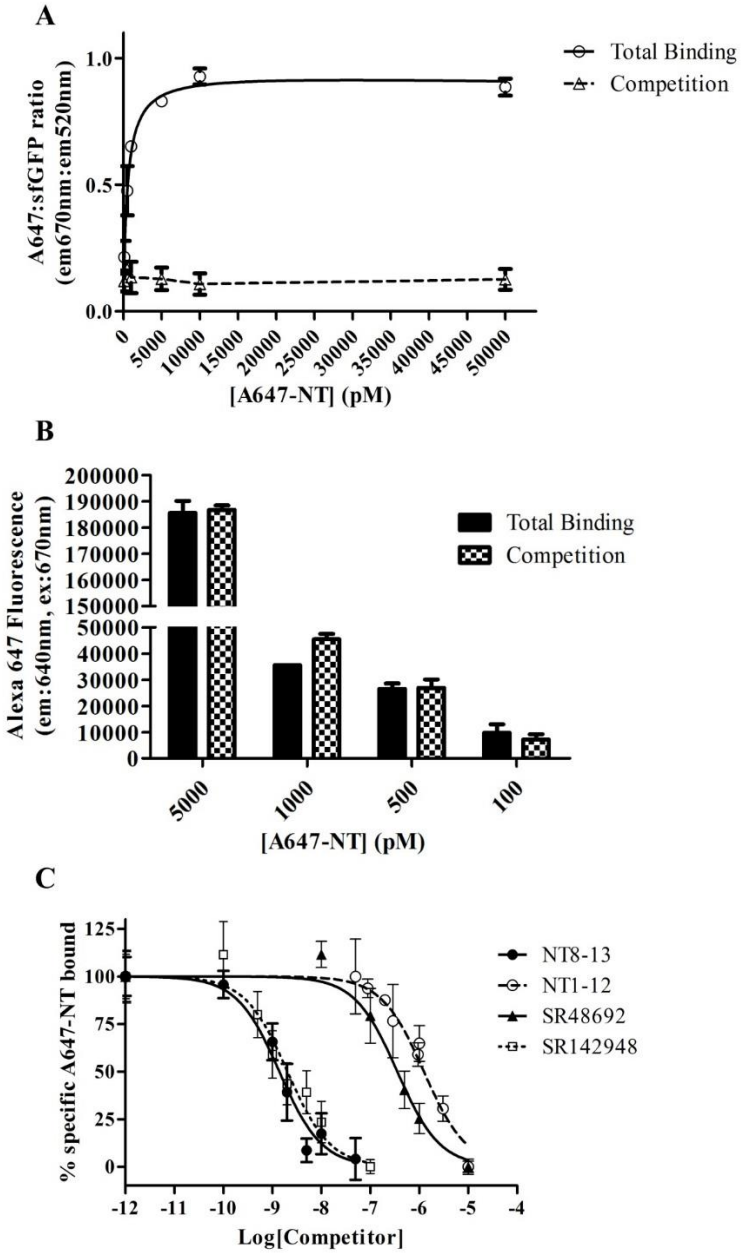


Figure 5.4 Saturation (A-B) and competition (C) binding assays were performed on solubilized, biotinylated and sfGFP tagged NTS1-H4, immobilized on streptavidin-coated magnetic beads. For saturation binding (A), beads were exposed to increasing concentrations of A647-NT in the presence of excess NT8-13 (COMPETITION, open triangles, dashed line) or without competition (TOTAL BINDING, open circles, solid line). The ratio of A647-NT to sfGFP fluorescence was measured to account for slight differences in bead concentrations across the 96-well plates. (B) Potential ligand depletion was assayed by measuring the levels of A647-NT retained in the binding wells from (A) upon removal of the receptor coated beads. A decrease in A647-NT in total binding wells (black columns) compared to the competition wells (chequered columns), where unlabelled NT would saturate the receptor, would indicate depletion of A647-NT in the binding step. No ligand depletion was observed. For competition binding (C), NTS1-H4 coated beads were exposed to 2 nM A647-NT and increasing concentrations of unlabelled NT8-13 (filled circles with solid line), NT1-12 (open circle with dashed line), SR48692 (filled triangles and solid line) and SR142948 (open squares with dotted line). The ratio of A647-NT to sfGFP fluorescence was measured, with the data sets normalised to 100% based on the fluorescence of wells containing no competitor. Mean values ± the standard error of the mean (SEM) are plotted from data pooled across 3 separate experiments.

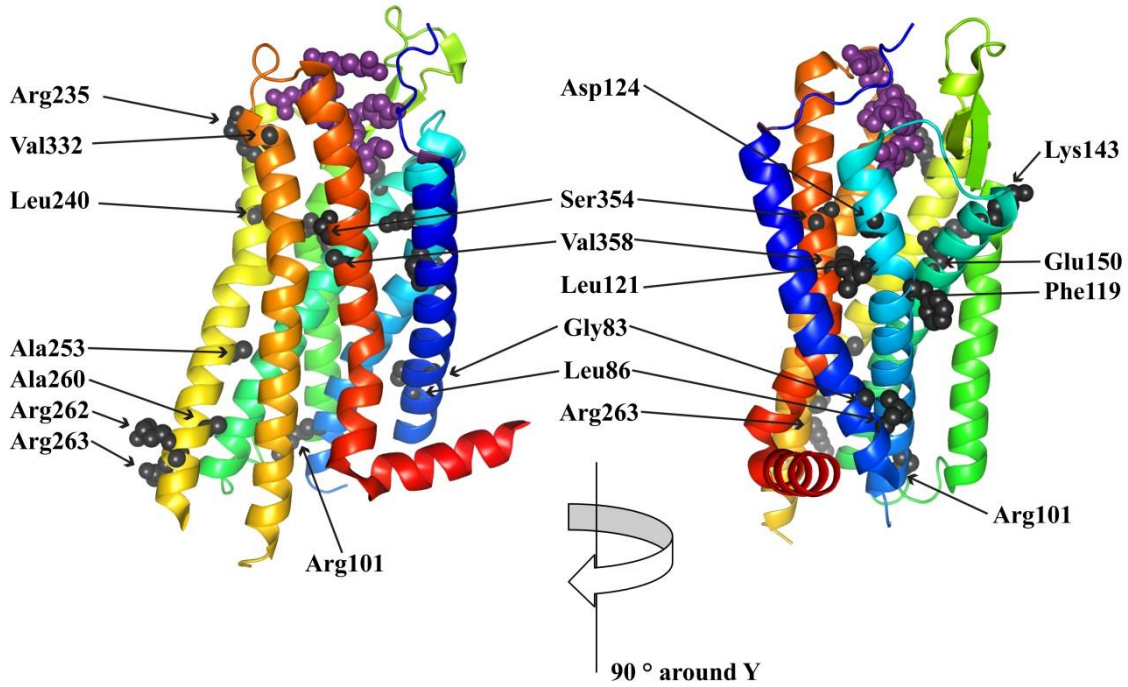


Figure 5.5 Positions of CHESS-selected mutations in the crystal structure of NTS1-H4 displayed at two angles rotated by 90 degrees. TM1 is coloured dark blue, TM2 light blue, TM3 cyan, TM4 green, TM5 yellow, TM6 orange, TM7 red and neurotensin purple.

5.1.6. References

[1] R. Fredriksson, H.B. Schioth, The repertoire of G-protein-coupled receptors in fully sequenced genomes, Mol. Pharmacol., 67 (2005) 1414-1425.

[2] J.P. Overington, B. Al-Lazikani, A.L. Hopkins, How many drug targets are there?, Nat. Rev. Drug. Discov., 5 (2006) 993-996.

[3] J.S. Mason, A. Bortolato, M. Congreve, F.H. Marshall, New insights from structural biology into the druggability of G protein-coupled receptors, Trends Pharmacol. Sci., (2012).

[4] K. Palczewski, T. Kumasaka, T. Hori, C.A. Behnke, H. Motoshima, B.A. Fox, I. Le Trong, D.C. Teller, T. Okada, R.E. Stenkamp, M. Yamamoto, M. Miyano, Crystal structure of rhodopsin: A G protein-coupled receptor, Science, 289 (2000) 739-745.

[5] V. Cherezov, D.M. Rosenbaum, M.A. Hanson, S.G. Rasmussen, F.S. Thian, T.S. Kobilka, H.J. Choi, P. Kuhn, W.I. Weis, B.K. Kobilka, R.C. Stevens, High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor, Science, 318 (2007) 1258-1265.

[6] V.P. Jaakola, M.T. Griffith, M.A. Hanson, V. Cherezov, E.Y. Chien, J.R. Lane, A.P. Ijzerman, R.C. Stevens, The 2.6 Ångström crystal structure of a human A2A adenosine receptor bound to an antagonist, Science, 322 (2008) 1211-1217.

[7] T. Warne, M.J. Serrano-Vega, J.G. Baker, R. Moukhametzianov, P.C. Edwards, R. Henderson, A.G. Leslie, C.G. Tate, G.F. Schertler, Structure of a beta1-adrenergic G-protein-coupled receptor, Nature, 454 (2008) 486-491.

[8] E.Y. Chien, W. Liu, Q. Zhao, V. Katritch, G.W. Han, M.A. Hanson, L. Shi, A.H. Newman, J.A. Javitch, V. Cherezov, R.C. Stevens, Structure of the human dopamine D3 receptor in complex with a D2/D3 selective antagonist, Science, 330 (2010) 1091-1095.

[9] B. Wu, E.Y. Chien, C.D. Mol, G. Fenalti, W. Liu, V. Katritch, R. Abagyan, A. Brooun, P. Wells, F.C. Bi, D.J. Hamel, P. Kuhn, T.M. Handel, V. Cherezov, R.C. Stevens, Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists, Science, 330 (2010) 1066-1071.

[10] T. Shimamura, M. Shiroishi, S. Weyand, H. Tsujimoto, G. Winter, V. Katritch, R. Abagyan, V. Cherezov, W. Liu, G.W. Han, T. Kobayashi, R.C. Stevens, S. Iwata, Structure of the human histamine H1 receptor complex with doxepin, Nature, 475 (2011) 65-70.

[11] K. Haga, A.C. Kruse, H. Asada, T. Yurugi-Kobayashi, M. Shiroishi, C. Zhang, W.I. Weis, T. Okada, B.K. Kobilka, T. Haga, T. Kobayashi, Structure of the human M2 muscarinic acetylcholine receptor bound to an antagonist, Nature, 482 (2012) 547-551.

[12] M.A. Hanson, C.B. Roth, E. Jo, M.T. Griffith, F.L. Scott, G. Reinhart, H. Desale, B. Clemons, S.M. Cahalan, S.C. Schuerer, M.G. Sanna, G.W. Han, P. Kuhn, H. Rosen, R.C. Stevens, Crystal structure of a lipid G protein-coupled receptor, Science, 335 (2012) 851-855.

[13] A. Manglik, A.C. Kruse, T.S. Kobilka, F.S. Thian, J.M. Mathiesen, R.K. Sunahara, L. Pardo, W.I. Weis, B.K. Kobilka, S. Granier, Crystal structure of the mu-opioid receptor bound to a morphinan antagonist, Nature, 485 (2012) 321-326.

[14] J.F. White, N. Noinaj, Y. Shibata, J. Love, B. Kloss, F. Xu, J. Gvozdenovic-Jeremic, P. Shah, J. Shiloach, C.G. Tate, R. Grisshammer, Structure of the agonist-bound neurotensin receptor, Nature, 490 (2012) 508-513.

[15] H. Wu, D. Wacker, M. Mileni, V. Katritch, G.W. Han, E. Vardy, W. Liu, A.A. Thompson, X.P. Huang, F.I. Carroll, S.W. Mascarella, R.B. Westkaemper, P.D. Mosier, B.L. Roth, V. Cherezov, R.C. Stevens, Structure of the human kappa-opioid receptor in complex with JDTic, Nature, 485 (2012) 327-332.

[16] C. Zhang, Y. Srinivasan, D.H. Arlow, J.J. Fung, D. Palmer, Y. Zheng, H.F. Green, A. Pandey, R.O. Dror, D.E. Shaw, W.I. Weis, S.R. Coughlin, B.K. Kobilka, High-resolution crystal structure of human protease-activated receptor 1, Nature, 492 (2012) 387-392.

[17] K. Hollenstein, J. Kean, A. Bortolato, R.K. Cheng, A.S. Dore, A. Jazayeri, R.M. Cooke, M. Weir, F.H. Marshall, Structure of class B GPCR corticotropin-releasing factor receptor 1, Nature, 499 (2013) 438-443.

[18] F.Y. Siu, M. He, C. de Graaf, G.W. Han, D. Yang, Z. Zhang, C. Zhou, Q. Xu, D. Wacker, J.S. Joseph, W. Liu, J. Lau, V. Cherezov, V. Katritch, M.W. Wang, R.C. Stevens, Structure of the human glucagon class B G-protein-coupled receptor, Nature, 499 (2013) 444-449.

[19] P. Egloff, M. Hillenbrand, C. Klenk, A. Batyuk, P. Heine, S. Balada, K. Schlinkmann, D.J. Scott, M. Schütz, A. Plückthun, Structure of signaling-competent neurotensin receptor 1 obtained by directed evolution in *E. coli*, Proc. Natl. Acad. Sci. U. S. A., 111 (2014) E655-662.

[20] E. Chun, A.A. Thompson, W. Liu, C.B. Roth, M.T. Griffith, V. Katritch, J. Kunken, F. Xu, V. Cherezov, M.A. Hanson, R.C. Stevens, Fusion partner toolchest for the stabilization and crystallization of G protein-coupled receptors, Structure, 20 (2012) 967-976.

[21] D.M. Rosenbaum, V. Cherezov, M.A. Hanson, S.G. Rasmussen, F.S. Thian, T.S. Kobilka, H.J. Choi, X.J. Yao, W.I. Weis, R.C. Stevens, B.K. Kobilka, GPCR engineering yields high-resolution structural insights into beta2-adrenergic receptor function, Science, 318 (2007) 1266-1273.

[22] J. Standfuss, G. Xie, P.C. Edwards, M. Burghammer, D.D. Oprian, G.F. Schertler, Crystal structure of a thermally stable rhodopsin mutant, J. Mol. Biol., 372 (2007) 1179-1188.

5.2. Abbreviations

2YT	2x concentrated yeast tryptone medium
AU	absorbance unit
CHAPS	3-[(3-cholamidopropyl)-dimethylammonio]-1-propane sulfonate
CD	circular dichroism (spectroscopy)
CHS	cholesteryl hemisuccinate
DDM	n-dodecyl-β-D-maltopyranoside
DM	n-decyl-β-D-maltopyranoside
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
DTT	dithiothreitol
EC	extracellular loop
E. coli	*Escherichia coli*
EDTA	ethylenediaminetetraacetic acid
FPLC	fast protein liquid chromatography
GdnHCl	guanidine hydrochloride
GPCR	G protein-coupled receptor
GRK	GPCR kinase
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HRV	human rhinovirus
IC	intracellular loop
IMAC	immobilized metal ion affinity chromatography
IPTG	isopropyl-β-D-thiogalactoside
K_D	echilibrium dissociation constant
kDa	kilo Dalton
k_{off}	dissociation rate constant
k_{on}	association rate constant
MALS	multi-angle light scattering
MBP	maltose binding protein
MR	molecular replacement
MSP	membrane scaffold protein

MW	molecular weight
NG	n-nonyl-β-D-glucopyranoside
NHS	N-hydroxysuccinimide
Ni-NTA	nickel-nitrilotriacetic acid
NM	n-nonyl-β-D-maltopyranoside
NMR	nuclear magnetic resonance (spectroscopy)
NT	neurotensin
NTR1	neurotensin receptor 1
OD	optical density
OG	n-octyl-β-D-glucopyranoside
OGNG	octyl glucose neopentyl glycol
OTM	n-octyl-β-D-thiomaltopyranoside
PCR	polymerase chain reaction
pD	protein D from phage lambda
PDB	Protein Data Bank
PDB-ID	Protein Data Bank identification number
PEG	polyethylene glycol
RMSD	root mean square deviation
RT	room temperature
SDS-PAGE	sodium dodecylsulfate polyacrylamide gel electrophoresis
SEC	size exclusion chromatography
SP	sulphopropyl
T_m	midpoint of transition
Tris	tris(hydroxymethyl)aminomethane
TROSY	transverse relaxation-optimized spectroscopy
TrxA	thioredoxin A

5.3. Plasmid for large-scale expression of NTR1 variants

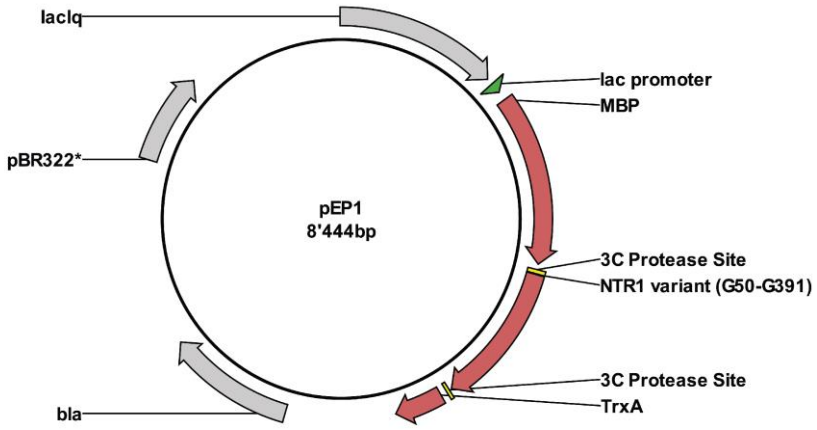


Figure 5.6 This vector was originally obtained as a kind gift from Dr. Reinhard Grisshammer (pRG) and subsequently adapted according to this illustration. As a result of directed evolution, it contains a mutation in the origin of replication (C7054G, indicated by an asterisk), which increased the copy number by approximately 100% compared to the parent plasmid[1]. The NTR1 variants were inserted in between the two 3C protease sites using the restriction enzymes BamH1 and Cfr9I. The fusion protein MBP encodes a periplasmic signal sequence at the N-terminus of the construct. A deca-histidine tag follows the C-terminal TrxA fusion, but it is not required for the purification procedure developed in this work.

5.4. Potential applicability of purification principles for other GPCRs

Table 5.2 List of class A GPCRs with potential for analogous, ligand-based receptor purification. Only receptors binding to non-modified peptides without lysines were included.

Receptor	Gene	Ligand	Affinity	Reference
Type-1 angiotensin II receptor (human)	AGTR1	DRVYIHPF	Ki ≈ 1.47 nM	[2]
		RVYIHPF	Ki ≈ 1.40 nM	[2]
Type-2 angiotensin II receptor[a]	AGTR2	DRVYIHPF	Ki < 1 nM	[3; 4]
		RVYIHPF	Ki < 10 nM	[3; 4]
B1 bradykinin receptor (human)	BDKRB1	RPPGFSPL	Ki ≈ 78 nM	[5; 6]
B1 bradykinin receptor (rat)	Bdkrb1	RPPGFSPF	Ki ≈ 15 nM	[3; 7]
		RPPGFSPL	Ki ≈ 31 nM	[3; 7]
		ISRPPGFSPF	Ki ≈ 46 nM	[3; 7]
B2 bradykinin receptor (human)	BDKRB2	RPPGFSPFR	IC50 ≈ 0.5 nM	[3; 8]
Neurotensin receptor 2 (human)	NTSR2	RRPYIL	Ki ≈ 1.4 nM	[9]
δ-type opioid receptor (human)	OPRD	YGGFL	Ki ≈ 2 nM	[3; 10]
μ-type opioid receptor (human)	OPRM1	YGGFL	Ki ≈ 7 nM	[3; 10]
κ-type opioid receptor (human)	OPRK1	YGGFLRRI	Ki ≈ 0.2 nM	[3; 10]

[a] species not defined in references

5.5. References

1. Sarkar, C. A., Dodevski, I., Kenig, M., Dudli, S., Mohr, A., Hermans, E. & Plückthun, A. (2008). Directed evolution of a G protein-coupled receptor for expression, stability, and binding selectivity. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 14808-14813.
2. Chiu, A. T., Dunscomb, J., Kosierowski, J., Burton, C. R. A., Santomenna, L. D., Corjay, M. H. & Benfield, P. (1993). The Ligand-Binding Signatures of the Rat at(1a), at(1b) and the Human at(1) Receptors Are Essentially Identical. *Biochem. Biophys. Res. Commun.* **197**, 440-449.
3. Okuno, Y., Yang, J. Y., Taneishi, K., Yabuuchi, H. & Tsujimoto, G. (2006). GLIDA: GPCR-ligand database for chemical genomic drug discovery. *Nucleic Acids Res.* **34**, D673-D677.
4. Degasparo, M., Husain, A., Alexander, W., Catt, K. J., Chiu, A. T., Drew, M., Goodfriend, T., Harding, J. W., Inagami, T. & Timmermans, P. B. M. W. M. (1995). Proposed Update of Angiotensin Receptor Nomenclature. *Hypertension* **25**, 924-927.
5. Bastian, S., Loillier, B., Paquet, J. L. & Pruneau, D. (1997). Stable expression of human kinin B-1 receptor in 293 cells: pharmacological and functional characterization. *Br. J. Pharmacol.* **122**, 393-399.
6. Austin, C. E., Faussner, A., Robinson, H. E., Chakravarty, S., Kyle, D. J., Bathon, J. M. & Proud, D. (1997). Stable expression of the human kinin B-1 receptor in Chinese hamster ovary cells - Characterization of ligand binding and effector pathways. *J. Biol. Chem.* **272**, 11420-11425.
7. Jones, C., Phillips, E., Davis, C., Arbuckle, J., Yaqoob, M., Burgess, G. M., Docherty, R. J., Webb, M., Bevan, S. J. & McIntyre, P. (1999). Molecular characterisation of cloned bradykinin B-1 receptors from rat and human. *Eur. J. Pharmacol.* **374**, 423-433.
8. Hess J. F., M. T., Stonesifer G. Y., Fraher J., Strader C. D., Ransom R. W. Ransom. (1993). Differential Pharmacology of Cloned Human and Mouse B2 Bradykinin Receptors. *Mol. Pharmacol.* **45**.
9. Richard, F., Barrosso, S., Martinez, J., Labbe-Jullie, C. & Kitabgi, P. (2001). Agonism, inverse agonism, and neutral antagonism at the constitutively active human neurotensin receptor 2. *Mol. Pharmacol.* **60**, 1392-1398.
10. Toll, L., Berzetei-Gurske, I. P., Polgar, W. E., Brandt, S. R., Adapa, I. D., Rodriguez, L., Schwartz, R. W., Haggart, D., O'Brien, A., White, A., Kennedy, J. M., Craymer, K., Farrington, L. & Auh, J. S. (1998). Standard binding and functional assays related to medications development division testing for potential cocaine and opiate narcotic treatment medications. *NIDA Res. Monogr.* **178**, 440-66.

5.6. Curriculum vitae

First Name	Pascal
Last Name	Egloff
Date of birth	September 14, 1983
Nationality	Swiss

Education

06.2009 – 2014	Dissertation in the Group of Prof. Dr. Andreas Plückthun, Institute of Biochemistry, University of Zurich, Switzerland. Title: Structures of Neurotensin Receptor 1 Variants – Native GPCRs Expressed in *E. coli*.
11.2013 – 12.2013	Guest Scientist in the Group of Prof. Dr. Gerhard Wagner, Harvard Medical School, Boston, USA.
08.2008 – 05.2009	Master's Thesis in the Group of Prof. Dr. William I. Weis, Stanford Medical School, Palo Alto, USA. Title: Structural and functional characterization of N-Cadherin-β-catenin interactions.
08.2007 – 06.2009	Master in biophysics and structural biology at ETH Zurich, Switzerland
08.2004 – 06.2007	Bachelor in biology at ETH Zurich, Switzerland
08.2002 – 07.2004	Kantonsschule Baden, Switzerland
08.1999 – 06.2002	Wirtschaftsdiplomschule Baden, Switzerland

Publications

Egloff P., De Luigi M., Heine P., Balada S. Plückthun A. (2014). A Cleavable Ligand Column for the Rapid Isolation of Large Quantities of Homogeneous and Functional Neurotensin Receptor 1 Variants from *E. coli*. Protein Expr. Purif. *Manuscript in preparation.*

Scott D. J., Kummer L., **Egloff P**., Bathgate R. A. D. and Plückthun A. (2014). Improving the apo-state detergent-stability of NTS1 with CHESS for pharmacological and structural studies. BBA Biomembranes. *Manuscript in preparation*

Egloff P., Hillenbrand M., Klenk C., Batyuk A., Heine P., Balada S., Schlinkmann K. M., Scott D.J., Schütz M., Plückthun A. (2014). Structure of signaling-competent neurotensin receptor 1 obtained by directed evolution in *E. coli*. Proc. Natl. Acad. Sci. USA, **111**, E655-62

Schubert M., Bleuler-Martinez S., Butschi A., Wälti M.A., **Egloff P**., Stutz K., Yan S., Collot M., Mallet J.M., Wilson I.B., Hengartner M.O., Aebi M., Allain F.H., Künzler M. (2012). Plasticity of the β-trefoil protein fold in the recognition and control of invertebrate predators and parasites by a fungal defence system. PLoS Pathog., **8**, E1002706

5.7. Acknowledgments

I would like to express my sincere gratitude to Prof. Dr. Andreas Plückthun for entrusting me with this challenging but fascinating scientific project and for providing an excellent research environment in all respects. I have appreciated the freedom that I had as a PhD student, which allowed me to implement my own ideas and to promote collaborations with numerous groups. Accordingly, this project would never have reached this level without the strong teams of scientists, all located either directly at the University of Zurich, in Boston at Harvard Medical School or at Utrecht University. In particular, I would like to thank Prof. Dr. Raimund Dutzler for very helpful advice regarding crystal handling, data collection and diffraction analysis. I thank Prof. Dr. Gerhard Wagner and Dr. Franz Hagn, who welcomed me warmly at HMS and who are excellent collaborators in the solution-state NMR project. Dr. Prof. Gert Folkers is pushing solid-state NMR on NTR1 and I thank him for his collaboration. My thesis committee supported this project with helpful counsel and included Prof. Dr. Raimund Dutzler, Prof. Dr. Stephan Grzesiek and Prof. Dr. Nenad Ban.

For the strong support in solving crystal structures and for the teaching of know-how in terms of structure determination, I would like to thank Dr. Alexander Batyuk and PD Dr. Peer Mittl. Matthias Hillenbrand and Dr. Christoph Klenk made significant contributions by characterizing signaling properties of evolved variants and by giving constructive feedback during many excellent discussions. A special thank goes to Dr. Karola Schlinkmann and to Dr. Daniel James Scott, who adjusted their invaluable directed evolution work to the needs of my project and generated the variants that I could crystallize. I would also like to thank Dr. Igor Dodevski, who supported my practical endeavors at the very beginning of this dissertation and even more important, he was one of the extremely rare persons, who seemed to believe already back then that I would be able to solve a GPCR structure – excellent mental support.

I am most grateful to Stephanie Balada, Mattia de Luigi, Philipp Heine and Christian Schori, who were all excellent and dedicated semester and master students. Their enthusiasm is exemplified by Philipp Heine and Stephanie Balada, who went on collecting valuable datasets in the middle of the night at the synchrotron in Villigen, when I was at my personal limits and could not go on anymore due to acute health problems and over tiredness. A big thank goes to all former and current members of the Plückthun group, but especially to Marco Schütz, Dr. Christian Reichen, Dr. Christoph Klenk, Mattia de Luigi, Jakob Stüber, Philipp Heine, Matthias Hillenbrand, Dr. Pablo Tomatis and Dr. Nikolas Stefan for their inputs and discussions but also for their friendship and many enjoyable hours outside the laboratories or inside the seminar room late at night while playing Rockband. It was great!

The BSM PhD-program is acknowledged as a great exchange platform for students during the annual retreats or during other social activities. For all the administrative support I obtained at the Institute of Biochemistry, I would like to thank Dr. Peter Lindner, Petra Vogt and Ilse Plückthun. For house-keeping services Maria Keremis-Meletiou and Bruno Bellofato, for the material supply Sibylle Strassmann, for computer maintenance Stephan Klauser and Steve Rast, for the excellent protein crystallization facility Beat Blattman and Céline Stutz-Ducommun, and for many kinds of additional support our outstanding workshop team Adrian Schmid and Sascha Weidner.

I would like to thank my former sifu Rinaldo di Lorenzo for helping me to develop a stronger will, which was clearly beneficial for my performance in this challenging project. I also acknowledge my former supervisor Dr. Prof. William I. Weis, whose brilliance and scientific objectivity are still inspiring me.

Finally I would like to thank the strong support of my parents, my sister and my friends. They motivated me with their interest in my work, but at the same time they helped me to stay in the world of common/normal people without getting lost in the molecular world. My deepest gratitude goes to Daniela for her love and patience and understanding.